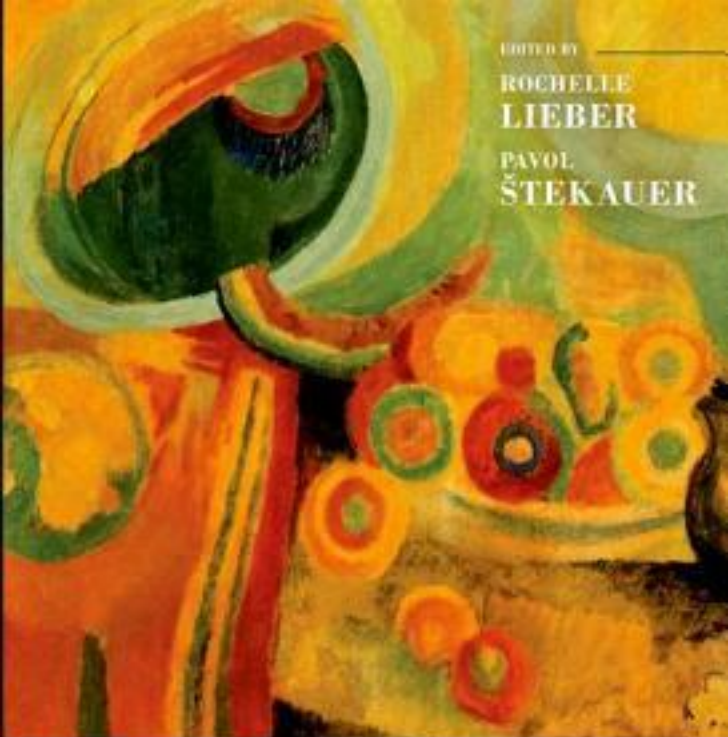


An abstract painting featuring a large, dark green, curved shape in the upper left, resembling a stylized head or a large letter 'C'. Below it, there are various organic, rounded forms in shades of yellow, orange, and red. Some of these shapes have concentric circles or smaller shapes inside them, creating a sense of depth and complexity. The background is a mix of these warm colors, with some areas appearing more textured or layered.

EDITED BY
ROCHELLE
LIEBER
PAVOL
ŠTEKAUER

≡ The Oxford Handbook of
**DERIVATIONAL
MORPHOLOGY**

THE OXFORD HANDBOOK OF

**DERIVATIONAL
MORPHOLOGY**

OXFORD HANDBOOKS IN LINGUISTICS

RECENTLY PUBLISHED

THE OXFORD HANDBOOK OF ARABIC LINGUISTICS

Edited by Jonathan Owens

THE OXFORD HANDBOOK OF COMPOSITIONALITY

Edited by Markus Werning, Wolfram Hinzen, and Edouard Machery

THE OXFORD HANDBOOK OF COMPOUNDING

Edited by Rochelle Lieber and Pavol Štekauer

THE OXFORD HANDBOOK OF CONSTRUCTION GRAMMAR

Edited by Thomas Hoffman and Graeme Trousdale

THE OXFORD HANDBOOK OF CORPUS PHONOLOGY

Edited by Jacques Durand, Ulrike Gut, and Gjert Kristoffersen

THE OXFORD HANDBOOK OF DERIVATIONAL MORPHOLOGY

Edited by Rochelle Lieber and Pavol Štekauer

THE OXFORD HANDBOOK OF GRAMMATICALIZATION

Edited by Heiko Narrog and Bernd Heine

THE OXFORD HANDBOOK OF HISTORICAL PHONOLOGY

Edited by Patrick Honeybone and Joseph Salmons

THE OXFORD HANDBOOK OF THE HISTORY OF ENGLISH

Edited by Terttu Nevalainen and Elizabeth Closs Traugott

THE OXFORD HANDBOOK OF THE HISTORY OF LINGUISTICS

Edited by Keith Allan

THE OXFORD HANDBOOK OF LABORATORY PHONOLOGY

Edited by Abigail C. Cohn, Cécile Fougeron, and Marie Hoffman

THE OXFORD HANDBOOK OF LANGUAGE AND LAW

Edited by Peter Tiersma and Lawrence M. Solan

THE OXFORD HANDBOOK OF LANGUAGE EVOLUTION

Edited by Maggie Tallerman and Kathleen Gibson

THE OXFORD HANDBOOK OF LINGUISTIC FIELDWORK

Edited by Nicholas Thieberger

THE OXFORD HANDBOOK OF SOCIOLINGUISTICS

Edited by Robert Bayley, Richard Cameron, and Ceil Lucas

THE OXFORD HANDBOOK OF TENSE AND ASPECT

Edited by Robert I. Binnick

[for a complete list of Oxford Handbooks in Linguistics please see pp. 928–929]

THE OXFORD HANDBOOK OF

DERIVATIONAL

MORPHOLOGY

Edited by
ROCHELLE LIEBER
and
PAVOL ŠTEKAUER

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© editorial matter and organization Rochelle Lieber and Pavol Štekauer 2014;
© the chapters their several authors 2014

The moral rights of the authors have been asserted

First Edition published in 2014

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2014938927

ISBN 978-0-19-964164-2

Printed and bound by
CPI Group (UK) Ltd, Croydon, CRO 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

CONTENTS

<i>List of Figures</i>	ix
<i>List of Tables</i>	xi
<i>Contributors</i>	xiii
<i>List of Abbreviations</i>	xxi

PART I

1. Introduction: The Scope of the Handbook ROCHELLE LIEBER AND PAVOL ŠTEKAUER	3
2. Delineating Derivation and Inflection PIUS TEN HACKEN	10
3. Delineating Derivation and Compounding SUSAN OLSEN	26
4. Theoretical Approaches to Derivation ROCHELLE LIEBER	50
5. Productivity, Blocking, and Lexicalization MARK ARONOFF AND MARK LINDSAY	67
6. Methodological Issues in Studying Derivation ROCHELLE LIEBER	84
7. Experimental and Psycholinguistic Approaches HARALD BAAZEN	95
8. Concatenative Derivation LAURIE BAUER	118
9. Infixation JULIETTE BLEVINS	136

10. Conversion	154
SALVADOR VALERA	
11. Non-concatenative Derivation: Reduplication	169
SHARON INKELAS	
12. Non-concatenative Derivation: Other Processes	190
STUART DAVIS AND NATSUKO TSUJIMURA	
13. Allomorphy	219
MARY PASTER	
14. Nominal Derivation	235
ARTEMIS ALEXIADOU	
15. Verbal Derivation	257
ANDREW KOONTZ-GARBODEN	
16. Adjectival and Adverbial Derivation	276
ANTONIO FÁBREGAS	
17. Evaluative Derivation	296
LÍVIA KÖRTVÉLYESSY	
18. Derivation and Function Words	317
GREGORY STUMP	
19. Polysemy in Derivation	338
FRANZ RAINER	
20. Derivational Paradigms	354
PAVOL ŠTEKAUER	
21. Affix Ordering in Derivation	370
PAULIINA SAARINEN AND JENNIFER HAY	
22. Derivation and Historical Change	384
CAROLA TRIPS	
23. Derivation in a Social Context	407
LÍVIA KÖRTVÉLYESSY AND PAVOL ŠTEKAUER	
24. Acquisition of Derivational Morphology	424
EVE V. CLARK	

PART II

25. Indo-European	443
PINGALI SAILAJA	
26. Uralic	473
FERENC KIEFER AND JOHANNA LAAKSO	
27. Altaic	493
IRINA NIKOLAEVA	
28. Yeniseian	509
EDWARD J. VAJDA	
29. Mon-Khmer	520
MARK J. ALVES	
30. Austronesian	545
ROBERT BLUST	
31. Niger-Congo	558
DENIS CREISSELS	
32. Afroasiatic	573
ERIN SHAY	
33. Nilo-Saharan	591
GERRIT J. DIMMENDAAL	
34. Sino-Tibetan	609
KAREN STEFFEN CHUNG, NATHAN W. HILL, AND JACKSON T.-S. SUN	
35. Pama-Nyungan	651
JANE SIMPSON	
36. Athabaskan	669
KEREN RICE	
37. Eskimo-Aleut	702
ALANA JOHNS	
38. Uto-Aztecan	724
GABRIELA CABALLERO	

39. Mataguyan	743
VERÓNICA NERCESIAN	
40. Areal Tendencies in Derivation	767
BERND HEINE	
41. Universals in Derivation	777
ROCHELLE LIEBER AND PAVOL ŠTEKAUER	
<i>References</i>	787
<i>Language Index</i>	885
<i>Name Index</i>	899
<i>Subject Index</i>	917

LIST OF FIGURES

4.1	Saussure's sign	51
4.2	Saussure's sign re-imagined for the sensory-motor system	51
4.3	Mapping in inflection	53
4.4	The simplex sign <i>cat</i>	57
4.5	The simplex sign <i>kill</i>	58
4.6a	Semantic composition in a lexical model	60
4.6b	Semantic composition in an inferential model	60
4.7a	Euthanize in an inferential model	61
4.7b	Euthanize in a lexical model	62
5.1	French borrowings as a percentage of all new words	77
5.2	Derived <i>-ity</i> as a percentage of all <i>-ity</i> words	78
5.3	Number of new English verbs and adjectives	79
5.4	New adjectives and verbs entering English, showing a rapid decline in the relative number of new verbs beginning in the 1600s	79
5.5	New derivations of <i>-ity</i> versus <i>-ment</i> over the past 750 years	80
11.1	Radial category for the semantics of reduplication	181
13.1	Types of allomorphy	225
17.1	The place of Evaluative morphology according to Scalise (1984)	299
17.2	An overview of categories by Bauer (2004b)	300
17.3	Synchronic-diachronic model of the semantics of diminutives	301
17.4	Model of evaluative word formation	306
17.5	Radial model of EM semantics	307
25.1	Indo-European groups	444
30.1	The Austronesian family tree	545
38.1	Uto-Aztecan language family	725

LIST OF TABLES

5.1	Comparing the productivity of <i>-ity</i> and <i>-ness</i>	74
5.2	<i>P*</i> value comparison	75
5.3	Derived forms for <i>-ment</i> and <i>-ity</i>	78
5.4	Sample Google ETM counts for high-frequency doublets	82
5.5	Sample Google ETM counts for high-frequency singletons	82
5.6	<i>-ical</i> is productive in stems ending in <i>olog</i>	82
9.1	Semantics of derivational infixes: from intangible to lexical	144
10.1	Different types of conversion and different types of languages. These results classify doubtful evidence as <i>Uncertain</i>	165
10.2	Word classes and different types of conversion. Only the types where a sufficient representation are shown. These results classify doubtful evidence as <i>Uncertain</i>	167
14.1	Formation of argument nouns in Saisiyat	241
17.1	Descriptive vs. qualitative perspective in evaluation	304
18.1	Categories of function words	318
18.2	Eight logically possible types of derivational relation involving function words	322
18.3	Deictic adverbs in Shughni	325
18.4	Demonstrative pronouns and deictic adverbs as derivational bases in Georgian	326
18.5	Interrogative, reflexive, indefinite, and personal pronouns as derivational bases in Georgian	326
18.6	Sanskrit derivative proforms in <i>-rhi</i> and <i>-ti</i>	327
18.7	Differences between the perfect auxiliary and the verb <i>have</i> 'own'	328
18.8	Some derived adverbs in Sanskrit	329
18.9	Some derived adjectives in Sanskrit	330
18.10	The distinction between ordinary nouns and pronouns in the <i>a</i> -stem and <i>ā</i> -stem declensions	331
18.11	Derived pronouns with comparative and superlative morphology in Sanskrit	332

18.12	Ordinal derivatives of compound cardinal numerals in five languages	334
18.13	Content words derived from numerals, mostly by conversion (American English)	336
24.1	Some innovative denominal verbs	429
24.2	Some innovative agent and instrument nouns	432
24.3	Using <i>un-</i> as a verbal prefix to talk about reversal	435
24.4	Innovative reversal verbs in French and German	435
25.1	Indo-European language family	445
25.2	Indo-Aryan—number of speakers	446
29.1	Core references on Mon-Khmer morphology	522
29.2	Common derivational processes in Mon-Khmer	523
29.3	Types of affixation in Mon-Khmer languages	525
29.4	Causative affixes in Mon-Khmer	526
29.5	Nominalizing affixes in Mon-Khmer	527
29.6	Demonstratives in Vietnamese	527
29.7	Specialized semantico-syntactic functions of affixation in Mon-Khmer	528
29.8	Number of days/years in Pacoh (Katuic)	534
29.9	Types of alternating reduplication with monosyllabic bases	536
29.10	Specialized semantico-syntactic functions of reduplication in Mon-Khmer	538
29.11	Types of lexical compounds in Mon-Khmer	541
30.1	Derivation by subtraction in the vocative forms of kinship terms	550
30.2	The reduplication-transitivity correlation in Tok Pisin	554
31.1	Valency change types and valency change markers in Wolof	563
34.1	Ideophone alternation patterns	647
35.1	Words for “boot” or “shoe” in some Pama-Nyungan languages	654
35.2	Words for “pig” in some Pama-Nyungan languages	657
35.3	Terms formed with “having” suffixes in some Pama-Nyungan languages	665
38.1	Valence stem allomorphy	734
38.2	Change of state predicates and thematic alternations	737
39.1	Derivational instrumental suffixes	753
39.2	Evaluative morphology suffixes	757

CONTRIBUTORS

Artemis Alexiadou is Professor of Theoretical and English Linguistics at the Universität Stuttgart. She received her Ph.D. in Linguistics in 1994 from the University of Potsdam. Her research interests lie in theoretical and comparative syntax, morphology, and most importantly in the interface between syntax, morphology, the lexicon, and interpretation. Her publications include books on the noun phrase (*Functional Structure in Nominals*, 2011, John Benjamins; *Noun Phrase in the Generative Perspective* together with Liliane Haegeman and Melita Stavrou, Mouton de Gruyter) as well as several journal articles and chapters in edited volumes on nominalization.

Mark J. Alves has been a professor in the Department of Reading, ESL, and Linguistics at Montgomery College in Rockville, Maryland, since 2004. His research, presentations, and publications have focused on historical, comparative, and typological linguistics in Southeast Asia with a concentration in Vietnamese and Mon-Khmer languages. Representative publications include “Distributional Properties of Mon-Khmer Causative Verbs” (2001), “The Vieto-Katuic Hypothesis: Lexical Evidence” (2005), *A Pacoh Grammar* (2006), “Pacoh Pronouns and Grammaticalization Clines” (2007), “Sino-Vietnamese Grammatical Vocabulary Sociolinguistic Conditions for Borrowing” (2009), among others.

Mark Aronoff is Distinguished Professor of Linguistics at Stony Brook University. His research touches on almost all aspects of morphology and its relations to phonology, syntax, semantics, and psycholinguistics. For the last dozen years he has been a member of a team studying a newly-created sign language, Al-Sayyid Bedouin Sign Language. From 1995 to 2001, he served as Editor of *Language*, the journal of the Linguistic Society of America.

Harald Baayen studied general linguistics at the Free University in Amsterdam. Following completion of his doctoral dissertation on morphological productivity with Geert Booij (linguistics) and Richard Gill (statistics) in 1989, he became a member of the research staff at the Max Planck Institute for Psycholinguistics in Nijmegen, where he focused his research on lexical processing. In 1998, he received a career advancement award from the Dutch research council NWO, which allowed him to strengthen his empirical research on morphological processing. In 2007, he took up a professorship in Edmonton, Canada, returning to Europe in 2011 to take up a chair in quantitative linguistics at the Eberhard Karls University in Tübingen, Germany, thanks to an Alexander von Humboldt research award. His current research focuses on discrimination learning

in language processing, computational modeling of lexical processing, articulatory, and statistical modeling of linguistic data with generalized additive mixed models.

Laurie Bauer is Professor of Linguistics and Dean of the Faculty of Research at Victoria University of Wellington, New Zealand. He is the author of several books on morphology, including *English Word-formation* (1983), *Introducing Linguistic Morphology* (1988, 2nd edition 2003), *Morphological Productivity* (2001), *A Glossary of Morphology* (2002), and, most recently, with Rochelle Lieber and Ingo Plag, *The Oxford Reference Guide to English Morphology* (2013). He was elected to a fellowship of the Royal Society of New Zealand in 2012.

Juliette Blevins is a professor of linguistics at the CUNY Graduate Center where she directs the Endangered Language Initiative. Her theory of Evolutionary Phonology (CUP, 2004) synthesizes work in sound change, phonetics, and typology, offering new explanations for a wide range of sound patterns and their distributions. Blevins has areal expertise in Austronesian, Australian Aboriginal, Native American, and Andamanese languages, and is currently working on the reconstruction of Proto-Ongan and Proto-Basque.

Robert Blust is Professor of Linguistics at the University of Hawai'i at Mānoa. He has conducted fieldwork on about 100 Austronesian languages, primarily in Borneo, Papua New Guinea, and Taiwan, and has authored over 220 publications, including the first single-authored book to cover the entire Austronesian language family (*The Austronesian Languages*, Pacific Linguistics, 2009). In addition, he has been working for years on the online *Austronesian Comparative Dictionary*, now at around 2,800 single-spaced printed pages.

Gabriela Caballero is Assistant Professor in the Department of Linguistics at the University of California, San Diego. Her main research focus concerns language documentation of endangered languages, the nature of intralinguistic and cross-linguistic variation in morphology and phonology, and languages of the Americas, especially Uto-Aztecan languages. She has recently published papers on the typology of Noun Incorporation, theoretical implications of the prosodic morphology of Guarijío, and topics in the phonology and morphology of Choguita Rarámuri, including affix order, multiple exponence, and morphological conditions on stress assignment.

Karen Steffen Chung (史嘉琳 Sh_Jialín), originally from St. Paul, Minnesota, USA, has taught English and linguistics in the Department of Foreign Languages and Literatures of National Taiwan University since 1990 and is currently Associate Professor. She gained her BA in East Asian Languages at the University of Minnesota in 1976; her MA in East Asian Studies at Princeton University in 1981; and her Ph.D. in Linguistics at the Universiteit Leiden in 2004, where her dissertation was entitled "Mandarin Compound Verbs".

Eve V. Clark is the Richard W. Lyman Professor in Humanities and Professor of Linguistics at Stanford University. She has done extensive cross-linguistic observational

and experimental research on children's semantic and pragmatic development, and on the acquisition of word formation. Her books include *The Ontogenesis of Meaning* (1979), *The Lexicon in Acquisition* (1993), and *First Language Acquisition* (2nd edn, 2009).

Denis Creissels retired in 2008 after teaching general linguistics at the universities of Grenoble (1971–96) and Lyon (1996–2008). His research interests center on linguistic diversity, the description of less-studied languages, and syntactic typology. He has been engaged in fieldwork on West African languages (Baule, Manding), Southern Bantu languages (Tswana), and Daghestanian languages (Akhvakh). His recent publications include descriptions of several Manding varieties (Kita Maninka, Mandinka, Niokolo Maninka). He is currently involved in projects on various Senegalese languages, including the edition of a volume on the noun class systems of Atlantic languages.

Stuart Davis is Professor of Linguistics at Indiana University, where he was chair of the linguistics department from 2004 to 2011. He has published extensively on issues of phonological analysis and theory, including matters arising from the phonology–morphology interface. While much of his work has a typological focus, he has published articles on such languages as American English, Arabic, Japanese, Korean, Italian, and Bambara.

Gerrit J. Dimmendaal is Professor of African Studies at the University of Cologne. In his research, he has focused on the Nilo-Saharan phylum, but he has also published on Niger-Congo and AfroAsiatic languages. His recent publications include an edited volume, *Coding Participant Marking: Construction Types in Twelve African Languages* (2009), and a course book, *Historical Linguistics and the Comparative Study of African Languages*. He is currently working on a reference grammar of a Niger-Congo language in Sudan, Tima, and a monograph on anthropological linguistics, *The Leopard's Spots: Essays on Language, Cognition and Culture*.

Antonio Fábregas is Full Professor of Hispanic Linguistics at the Language and Linguistics institute in the University of Tromsø, and affiliate to CASTL. His work has concentrated on neoconstructionist approaches to word formation and the lexicon. Among his most cited publications there are *The Internal Syntactic Structure of Relational Adjectives* (2007), *A Syntactic Account of Affix Rivalry in Spanish Nominalisations* (2010), and *Evidence for Multidominance in Spanish Agentive Nominalizations* (2012).

Pius ten Hacken is Universitätsprofessor at the Institut für Translationswissenschaft of the Leopold-Franzens-Universität Innsbruck. Formerly he was at Swansea University. His research interests include morphology, terminology, and the philosophy and history of linguistics. He is the author of *Defining Morphology* (Olms, 1994) and of *Chomskyan Linguistics and its Competitors* (Equinox, 2007), the editor of *Terminology, Computing and Translation* (Narr, 2006), and co-editor of *The Semantics of Word Formation and Lexicalization* (Edinburgh University Press, 2013).

Jennifer Hay is Associate Professor of Linguistics at the University of Canterbury, Christchurch, New Zealand, and a member of the New Zealand Institute of Language,

Brain & Behavior. Her fields of research include morphology, phonetics, sociolinguistics, laboratory phonology, sociophonetics, and New Zealand English. She is the author of *Causes and Consequences of Word Structure* (Routledge, 2003), and co-author of *Probabilistic Linguistics* (MIT Press, 2003), *New Zealand English: Its Origins and Evolution* (CUP, 2004), and *New Zealand English* (Edinburgh University Press, in press).

Bernd Heine is Emeritus Professor at the Institut für Afrikanistik, University of Cologne. He has held visiting professorships in Europe, Eastern Asia (Japan, Korea, China), Australia, Africa (Kenya, South Africa), North America (University of New Mexico, Dartmouth College), and South America (Brazil). His 33 books include *Possession: Cognitive Sources, Forces, and Grammaticalization* (CUP, 1997); *Auxiliaries: Cognitive Forces and Grammaticalization* (OUP, 1993); *Cognitive Foundations of Grammar* (OUP, 1997) (with Tania Kuteva); *World Lexicon of Grammaticalization* (CUP, 2002); *Language Contact and Grammatical Change* (CUP, 2005); *The Changing Languages of Europe* (OUP, 2006), and *The Evolution of Grammar* (OUP, 2007); and with Heiko Narrog as co-editor *The Oxford Handbook of Linguistic Analysis* (OUP, 2011) and *The Oxford Handbook of Grammaticalization* (OUP, 2012).

Nathan W. Hill is Lecturer in Tibetan and Linguistics at SOAS, University of London. Educated at Harvard University, his research focuses on Tibetan historical grammar and Tibeto-Burman comparative linguistics. He is the author of *A Lexicon of Tibetan Verb Stems as Reported by the Grammatical Tradition* (Munich, 2010) in addition to more than 25 articles.

Sharon Inkelas is Professor in the Department of Linguistics at the University of California, Berkeley, where she has taught since 1992. Inkelas received her Ph.D. from Stanford University in 1989 and has also held positions at UCLA and the University of Maryland. Her research focuses on the phonology–morphology interface; in 2005 she published *Reduplication: Doubling in Morphology*, co-authored with Cheryl Zoll.

Alana Johns teaches Linguistics at the University of Toronto, where she specializes in morphology and syntax. For over 20 years she has been researching morphosyntactic properties of the Inuit language, including dialects spoken in Nunatsiavut (Labrador), Iqaluit, and Qamani'tuaq (Baker Lake). She has published on ergativity (e.g. Deriving ergativity, 1992, *Linguistic Inquiry*), noun incorporation (e.g. Restricting noun incorporation: root movement, 2007, *Natural Language and Linguistic Theory*), and dialect differences (e.g. Eskimo-Aleut languages, 2010, *Language and Linguistics Compass*). She also works with community language specialists who are involved in language maintenance and/or language research.

Ferenc Kiefer was born on May 24, 1931, in Apatin. He studied mathematics, German and French linguistics. From 1973 until his retirement in 2001 he was Research Professor at the Research Institute for Linguistics of the Hungarian Academy of Sciences. His research interests include morphology, semantics (especially lexical semantics), and pragmatics (especially the semantics–pragmatics interface). He is a member of several

learned societies and academies (Hungarian Academy of Sciences (1987), Academia Europaea (1993), Austrian Academy of Sciences (1995), Honorary Member of the Linguistic Society of America (1996), Honorary Member of the Philological Society of Great Britain (1998)). He received an honorary doctorate from the University of Stockholm (1992), from the Université de Paris 13 (2001), and from the University of Szeged (2006).

Andrew Koontz-Garboden (Ph.D. 2007, Stanford University) is Senior Lecturer in Linguistics in the Department of Linguistics and English Language at the University of Manchester. His expertise lies in the cross-linguistic study of the lexical semantics/morphosyntax interface. He has published on issues in this area in, among other journals, *International Journal of American Linguistics*, *Natural Language and Linguistic Theory*, *Natural Language Semantics*, *Linguistic Inquiry*, and *Linguistics and Philosophy*.

Lívia Körtvélyessy graduated in English and German philology in 1996. She was awarded her Ph.D. at the Slovak Academy of Sciences in 2008. In the same year she became a member of the Department of British and American Studies at Pavol Jozef Šafárik University, Košice. Her fields of expertise are evaluative morphology, word formation, and linguistic typology. She is author of a monograph (published in Slovak) on the influence of sociolinguistic factors on word formation, and is co-editor (with Nicola Grandi) of *Handbook of Evaluative Morphology* (forthcoming in 2014 from Edinburgh University Press).

Johanna Laakso, born 1962, studied Finnic languages, Finno-Ugric, and general linguistics at the University of Helsinki and defended her Ph.D. thesis at the University of Helsinki in 1990. Since 2000 she holds the chair of Finno-Ugric language studies at the University of Vienna. Her main research interests include historical and comparative Finno-Ugric linguistics, morphology (in particular, word formation), contact linguistics and multilingualism, and gender linguistics.

Rochelle Lieber is Professor of Linguistics at the University of New Hampshire. Her interests include morphological theory, especially derivation and compounding, lexical semantics, and the morphology–syntax interface. She is the author of several books: *On the Organization of the Lexicon* (IULC, 1981), *An Integrated Theory of Autosegmental Processes* (State University of New York Press, 1987), *Deconstructing Morphology* (University of Chicago Press, 1992), *Morphology and Lexical Semantics* (CUP, 2004), and *Introducing Morphology* (CUP, 2010). She is co-author, with Laurie Bauer and Ingo Plag of *The Oxford Reference Guide to English Morphology* (OUP, 2013). Together with Pavol Štekauer she has edited two handbooks, *The Handbook of Word Formation* (Springer, 2005) and *The Oxford Handbook of Compounding* (OUP, 2009).

Mark Lindsay earned his Ph.D. from Stony Brook University. His dissertation research focused on exploring productivity and self-organization in the lexicon using corpora and evolutionary modeling. His published work has dealt with gathering and analyzing suffix productivity using the World Wide Web and dictionaries, as well as pop culture

linguistic phenomena, such as American English *iz*-infixation and the German *Inflektiv* (or *Erikativ*).

Verónica Nercesian is a researcher at CONICET (National Council of Scientific and Technical Research) and the Linguistic Research Institute, National University of Formosa. She teaches Linguistics and Lexical Theory at the National University of Buenos Aires. Her current interests include Wichi dialectal variation and verbal art, and the linguistic level interplay. Her Ph.D. thesis focused on Wichi grammar and the interplay of phonology, morphology, syntax, and semantics in word formation.

Irina Nikolaeva is Professor of Linguistics at SOAS (University of London). She has studied in Moscow and San Diego and received a Ph.D. in Linguistics from the University of Leiden in 1998. Her interests lie in the field of linguistic typology, lexicalist theories of grammar, and documentation and description of endangered languages. She has published several books on Uralic, Altaic, and Palaeosiberian languages based on extensive fieldwork, as well as works on syntax, semantics, information structure, and historical-comparative linguistics.

Susan Olsen received her Ph.D. in German and English linguistics at the University of Cologne, Germany, and earned tenure in the Department of Germanic Studies at Indiana University. She has held professorships at the University of Stuttgart and Leipzig. Since 2002 she has been Professor of English Linguistics at the Humboldt University in Berlin. Her publications include topics in syntax, lexical semantics, word formation, morphology, and the lexicon.

Mary Paster (BA, Ohio State University; MA and Ph.D., University of California, Berkeley) is Associate Professor and Chair of Linguistics and Cognitive Science at Pomona College in Claremont, California. She specializes in phonology and morphology and their interfaces, particularly in the study of tone systems, allomorphy, and affix ordering. Her research focuses on underdescribed African languages.

Franz Rainer is a full professor and Director of the Institute for Romance Languages at WU Wirtschaftsuniversität Wien), the Vienna University of Economics and Business. He received his first degree and doctorate in Romance languages and linguistics from the University of Salzburg, completed his Habilitation at the same institution in 1992, and was appointed to a chair at WU in 1993. He has been a corresponding member of the Austrian Academy of Sciences since 2000, and a full member since 2010. That same year he was also elected a member of the Academy of Europe (Academia Europaea). His main research interest lies in the area of word formation. He is an author of *Spanische Wortbildungslehre* (Niemeyer, 1993); *Carmens Erwerb der deutschen Wortbildung* (Verlag der Österreichischen Akademie der Wissenschaften, 2010); and co-edited (with M. Grossmann) the volume *La formazione delle parole in italiano* (Niemeyer, 2004).

Keren Rice is University Professor at the University of Toronto. She has studied Athabaskan languages for many years, and is author of *A Grammar of Slave* (Mouton

de Gruyter), which received the Bloomfield Book Award from the Linguistic Society of America. She has published many articles on Athabaskan languages as well as on topics in phonology. She is the author the book *Morpheme Order and Semantic Scope: Word Formation in the Athapaskan Verb* (CUP), and is co-editor of several books on Athabaskan languages. She serves as editor of the *International Journal of American Linguistics*.

Pauliina Saarinen is a Ph.D. candidate in Linguistics at the University of Canterbury in Christchurch, New Zealand. She is also affiliated with the New Zealand Institute of Language, Brain and Behavior (NZILBB), a multi-disciplinary research institute located at the University of Canterbury. Pauliina's Ph.D. research focuses on the production and perception of consonant duration in Finnish morphological paradigms.

Pingali Sailaja is Professor of English in the Centre for English Language Studies, University of Hyderabad, India. Her interests are in the areas of phonology and morphology, varieties of English, historical and linguistic aspects of English in India, and the teaching of English as a second language. Her books include *English Words: Structure, Formation and Literature* (2004) and *Indian English* (2009).

Erin Shay is an adjunct assistant professor of linguistics at the University of Colorado, Boulder. She is the author of three Chadic grammars and the author or co-author of numerous books, chapters, and papers on Chadic and descriptive and comparative linguistics. Her research has been funded by grants from the NSF, NEH, ACLS, and other institutions.

Jane Simpson studies the structure, use, and history of several Pama-Nyungan languages: Warumungu, Kaurana, and Warlpiri. She has worked on language maintenance, including producing *A Learner's Guide to Warumungu: Mirlamirlajinjiki Warumunguku apparrka* (IAD Press, 2002). Three current projects are a longitudinal study of Aboriginal children acquiring creoles, English and traditional languages, and a study of kinship and social categories in Australia. She is Chair of Indigenous Linguistics at the Australian National University.

Pavol Štekauer is Professor of English linguistics at P. J. Šafárik University, Košice, Slovakia. His research has focused on an onomasiological approach to word formation, sociolinguistic aspects of word formation, meaning predictability of complex words, and cross-linguistic research into word formation. He is the author of *A Theory of Conversion in English* (Peter Lang, 1996), *An Onomasiological Theory of English Word-Formation* (John Benjamins, 1998), *English Word-Formation: A History of Research (1960–1995)* (Gunter Narr, 2000), and *Meaning Predictability in Word-Formation* (John Benjamins). He co-edited (with Rochelle Lieber) *Handbook of Word-formation* (Springer 2005) and *Oxford Handbook of Compounding* (OUP, 2009).

Gregory Stump is Professor of Linguistics at the University of Kentucky. He has written extensively on a range of morphological topics, most of them relating to the structure and typology of inflectional systems. He is the author of *Inflectional Morphology* (CUP,

2001) and (with Raphael A. Finkel) of *Morphological Typology* (CUP, 2013); he also serves as co-editor of the journal *Word Structure*.

Jackson T.-S. Sun is a research fellow at the Institute of Linguistics in Academia Sinica (Taiwan). The focus of his research is on synchronic and diachronic phonology and morphosyntax of Bodic, Tani, and Qiangic languages in the Sino-Tibetan family. His major contributions comprise a monograph on Amdo Tibetan phonology (1996), phonological reconstruction of Proto-Tani (1993), proposal of the Rgyalrongic languages as a distinct subgroup (2000), and discovery of a new secondary articulation type “uvularization” in Qiang and neighboring languages (2013).

Carola Trips is Professor of English Linguistics at the University of Mannheim. She received her Ph.D. from the University of Stuttgart in 2001. Her main research interests have been diachronic syntax and morphology, lexical semantics, and linguistic theory. She is the author of a number of articles on these topics and of the following books: *From OV to VO in Early Middle English* (John Benjamins, 2002), *Diachronic Clues to Synchronic Grammar* (edited with Eric Fuß, John Benjamins, 2004), and *Lexical Semantics and Diachronic Morphology: The development of -hood, -dom, and -ship in the history of English* (Niemeyer, 2009).

Natsuko Tsujimura is Professor and Chair of the Department of East Asian Languages and Cultures at Indiana University and has been review editor for *Language*. She has published widely on almost all areas of Japanese linguistics with a particular research focus on lexical semantics. She is the author of *An Introduction to Japanese Linguistics* (Wiley-Blackwell, 3rd edition, 2013) and editor of *The Handbook of Japanese Linguistics* (Blackwell, 1999).

Edward J. Vajda is Professor of Russian language and culture, linguistics, and Inner and North Asian peoples in the Modern and Classical Languages of Western Washington University. He directs the Linguistics Program and is involved in documenting Ket, an endangered language of Siberia spoken by fewer than 50 people near the Yenisei River. He is the author of *Subordination and Coordination Strategies in North Asian Languages* (John Benjamins, 2008), *Languages and Prehistory of Central Siberia* (John Benjamins, 2004), and a number of articles devoted to Ket and other languages of Siberia.

Salvador Valera was born in Jaén, Spain, in 1967, graduated in English Philology from the University of Granada in 1990, and was awarded his Ph.D. in 1994 for a dissertation on the formal identity between English adjectives and adverbs. He is currently Senior Lecturer (tenured) at the University of Granada. His major interests are corpus linguistics and English morphology and syntax.

LIST OF ABBREVIATIONS

1	first person
2	second person
3	third person
ABL	ablative
ABS	absolutive
ACC	accusative
ACT	active
Adj	adjective
Adv	adverb
AFF	affirmative
AG	agent
AGR	agreement
AI	animate intransitive verb stem
AL	alienable possession
ALL	allative
AN	Austronesian
AN	animate
ANTIP	antipassive
AOR	aorist
AP	antipassive
APPL	applicative
ART	article
AS	Aslian
ATTR	attributive
AUG	augment
AUG	augmentative
AUGM	augmentative
AUX	auxiliary
AV	actor voice
BN	Bahnaric
CAUS	causative
CEMP	Central-Eastern Malayo-Polynesian
CER	certainty
CF	centrifugal

CF	circumfix
Cho	Chorote
CL	noun class marker
CL	classifier
CLX	noun class marker of class X
CM	conjugation marker
CMP	Central Malayo-Polynesian
COM	comitative
COMIT	comitative
COMP	comparative
COMPL	completive
COND	conditional
CONJ	conjunction
CONT	continuous aspect
CONV	converb
CP	centripedal
D	declarative
D.PAST	distant past
DAT	dative
DEF	definite
DEM	demonstrative
DETR	detransitive
DETRANS	detransitive
DIM	diminutive
DIR	directional
DIRV	directive
DIST	distal
DIST	distributive
DM	discourse marker
DN	deverbal noun
DU	dual
DUB	dubitative
DUR	durative
EMP	Eastern Malayo-Polynesian
EMPH	emphatic
ep	epenthetic
ER	evaluative rule
ERG	ergative
EUPH	euphonious
EXCL	exclusive
F	feminine
F	Formosan

FCT	factive
FEM	feminine
FOC	focalization
FOC	focus
FUT	future
FV	final vowel
GEN	genitive
GO	goal
GP	generic person
H	head-marking
HAB	habitual
HTR	high transitivity
IC	incorporation closer
IDEO	ideophone
IDN	identifier
IMP	imperative
IMPV	imperfective
INACC	inaccusatif (unaccustative)
INAN	inanimate
INC	inceptive
INCH	inchoative
INCL	inclusive
IND	indicative
INDEF	indefinite
INDIC	indicative
INF	infinitive
INFL	inflection
INS	instrumental
INST	instrumental
INSTR	instrument
INTERR	interrogative
INTR	intransitive
INTRANS	Intransitive
INV	inverse
IT	itive
KR	Khmeric
KS	Khasic
KT	Katuic
KU	Khmuic
LIG	ligature
LNK	attributive linker
LNK	linking element

LOC	locative
LTR	low transitivity
LV	locative voice
M	masculine
Ma	Maka
MAN	manner
MAS/MASC	masculine
ME	Middle English
MED	meditative evidential
MG	Mangic
MI	middle voice
MID	middle
MN	Monic
MP	Malayo-Polynesian
N	noun
NC	Nicobarese
NCM	noun class marker
ND	Nyangumarta dictionary
NEG	negation
NEUT	neuter
NFUT	non-future
NHG	New High German
Ni	Nivacle
NMLZ	nominalizer
NOM	nominalizer
NOM	nominative
NOMZ	nominalizer
NPAST	non-past
OBJ	object
OBL	oblique
OC	Oceanic
OE	Old English
OF	Old French
PART	participle
PASS	passive
PDE	Present-day English
PERF	perfect
PFV	perfective
PL	plural/pluractional
PL	Palaungic
POS/POSS	possessive
POSS	possessor

POT	potential
PRED	predicate
PREP	preposition
PRES	present (indicative)
PRF	perfect
PRIV	privative
PRO	pronominal
PROG	progressive
PROPR	propriative
PROX	proximal
PRS	present
PST	past
PTCP	participial
PTCP	participle
PURP	purposive
PV	patient voice
PV.PERF	patient voice perfective
QUOT	quotative
RAPPR	rapproachant (approaching)
R.PAST	recent past
RDP	reduplication
RECP	reciprocal
RED	reduplication
REDUP	reduplication
REFL	reflexive
REL	relativizer
REV	reversive
RFL	reflexive
RL	relational noun
S	subject
S	singular
SA	unglossable particle
SBJ	subject
SG	singular
SHWNG	South Halmahera-West New Guinea
SING	singular
SOC	sociative
Sp	Spanish
SPON	spontaneous
SUB	subordinator
SUB	subject
SUBJ	subjunctive

SUBS	subsecutive
SUFF	suffix
THEM	thematic morpheme
TOP	topic
TR	transitivizer
TRANS	transitive
TRN	transnumeral
TRS	transitivizer
UNDF	undefined
V	verb
VBLZ	verbalizer
VBZ	verbalizer
VEN	ventive
VN	Vietnamese
VT	Vietic
VWF	word-formation value
Wi	Wichi
WMP	Western Malayo-Polynesian
YY	Yir Yoront

PART I

CHAPTER 1

INTRODUCTION

The Scope of the Handbook

ROCHELLE LIEBER AND PAVOL ŠTEKAUER

1.1 WHY DERIVATION ON ITS OWN?

THIS handbook is intended as a companion to our earlier *Oxford Handbook of Compounding* (2009), as well as to the *Oxford Handbook of Inflection* (Baerman in press), and the *Oxford Handbook of Morphological Theory* (Audring and Masini forthcoming). We might justify it simply on the basis of symmetry, as part of an effort to cover all areas of the study of morphology thoroughly in this series. Nevertheless, we ought to have a better reason in mind for compiling a book of this sort. In this Handbook we hope to look at derivational morphology on its own terms to see what is distinctive about it, what defines it as a phenomenon, and how it is manifested in the languages of the world.

What do we mean by “derivation on its own terms”? To determine this, we must start first with defining what we mean by word formation. The term “word formation” refers to the creation of new lexemes in a language and is generally said to be composed of compounding and derivation. By “derivation” we therefore mean to refer to those parts of word formation other than compounding, a definition that is also used by Aikhenvald (2007: 1). Although the definition of “compounding” is by no means straightforward, we have dealt with it extensively in our Introduction to the *Oxford Handbook of Compounding*. For our purposes here, it is sufficient to make use of Bauer’s (2003: 40) definition of a compound as “the formation of a new lexeme by adjoining two or more lexemes.”¹ What we are left with when we subtract compounding from word formation are ways of creating new lexemes other than putting two or more lexemes together. In formal terms, this encompasses various kinds of affixation (prefixation, suffixation, infixation, circumfixation), but also

¹ We remain neutral on whether noun-incorporation is to be treated as a sort of compounding or as a matter of syntax. We assume, however, that it is not to be included as a part of derivation.

reduplication, templatic or root and pattern word formation, subtractive word formation, conversion, and miscellaneous tone and stress changing operations, specifically when they are not used for the purposes of inflection.

Approached from the perspective of function or semantics, we might define the core of derivation as including, but not limited to the creation of:

- event, process, and result lexemes;
- personals, including agent and patient;
- lexemes expressing non-inflectional gender (e.g. *actress*);
- lexemes expressing location in time and space;
- instruments;
- collectives and abstracts;
- evaluatives (including both size and attitude);
- negatives and privatives;
- lexemes relating to non-evaluative size and quantity;
- causatives, anti-causatives, applicatives, factitives, inchoatives, duratives, and the like.

Derivation may be either category-changing, or non-category-changing; for example, personal nouns may be formed from verbs (*writer*, *accountant*) but also from other nouns (*Londoner*, *pianist*). Verbs can be created from nouns or adjectives (*unionize*, *civilize*), or can be formed from other verbs, such as the causatives and applicatives that are typical of the Niger-Congo languages (Creissels, this volume). There are no doubt many other semantic categories into which derivation can fall, especially if we take into account the sort of lexical derivation that is to be found in polysynthetic languages, such as those of the Athabaskan (Rice, this volume) or Eskimo-Aleut languages (Johns, this volume). Indeed, some semantic categories can be quite idiosyncratic, as is the case with the suffix *-ier* in French, which creates names of trees from names of the respective fruit (*poire* ‘pear’ ~ *poirier* ‘pear tree’).

It would be convenient, of course, if we could take the intersection of these formal and functional categories and be left with a clearly delineated domain of derivation as the subject of this handbook. But language is rarely so obliging and we must acknowledge that on all sides we are faced with fuzzy boundaries. In some cases there is difficulty separating derivation from compounding. As Olsen (this volume) points out, identifying the point at which an independent lexeme becomes an affix is almost impossible to do. Or consider the case of reduplication. Some authors (e.g. Štekauer et al. 2012) treat full reduplication as a form of compounding, apart from partial reduplication; there is something to be said for this choice, as full reduplication certainly does fulfill the main criterion of compounding as being the composition of two lexemes. Still, others (Inkelas, this volume) find the most salient characteristic of reduplication—repetition—sufficient to treat full reduplication as a phenomenon distinct from compounding. On the other side, there are cases where the boundaries between derivation and inflection are indistinct, as with evaluatives in languages that have extensive noun class systems,

with certain classes being reserved for diminutives or augmentatives (see Creissels, this volume). Indeed, the puzzling nature of evaluatives has led some researchers to treat it as distinct from either derivation or inflection (see Körtvélyessy, this volume).

In spite of difficulties of this sort, the present volume is predicated on the assumption that there is something in the intersection between the formal means and the functional/semantic territory covered by derivation that defines a coherent field of study. Is this the case? Oddly, this is a question that does not seem to have been asked. One reason for this is that derivation has only rarely been treated apart from other sorts of morphology—compounding on the one hand and inflection on the other.

1.2 A BRIEF FORAY INTO HISTORY

We do not mean to dwell on the historical development of the field of morphology, as this is a subject that has already been covered in our *Handbook of Word Formation* (2005) and is to be the subject of *The Oxford Handbook of Morphological Theory* (Audring and Masini forthcoming). But at least a brief mention of the treatment of derivation in morphological theory seems justified here. Seminal works in the American structuralist tradition, such as Harris (1946) or Hockett (1947, 1954) were preoccupied with methods of analyzing morphemes, and do not seem to provide separate treatments of inflection and derivation.² Nor do some of the key works in morphology from the middle of the 20th century single out derivation as a distinct matter for study. Lees' *The Grammar of English Nominalizations* (1960) represents early work characteristic of the generative tradition in North America. Lees focuses primarily on noun-noun compounds, but also assumes that transformations of various sorts can introduce category-changing derivational morphology, in particular affixes that nominalize verbs in English. Marchand's *The Categories and Types of Present-day English Word-Formation* (1960/9) is representative of the mid-century view on word formation in Western Europe. The scope of Marchand's work, drawing mainly on the structuralist tradition of the Geneva School and the ideas of Coseriu (1952), is much broader, covering a wide range of word-formation processes in English derivation and compounding. Dokulil's *Tvoření slov v češtině I. Teorie odvozování slov* [Word-Formation in Czech. A Theory of Word Derivation] (1962) is representative of the field in Central Europe. His is the most comprehensive theory from among the authors of the 1960s.³ Dokulil discusses and foreshadows a number of topics which have become central to the field of derivational morphology, including a general onomasiological theory of word formation, individual word-formation processes and cognitive foundations of these processes, productivity,

² Bloomfield (1933: 237) indeed implies that the distinction between inflection and what we would call derivation does not necessarily apply in all languages.

³ Unfortunately, his publications were not written in English, so they have had limited influence in North America or Western Europe.

derivational paradigms, and lexicalization, among others. His work continues to be of influence among morphologists in Central Europe.

Subsequent work has only rarely singled out derivation from compounding and inflection. Indeed, Aronoff's *Word Formation in Generative Grammar* (1976) seems to be the lone example.⁴ Aronoff is careful to distinguish derivation from inflection, the latter being a matter of syntax: he mentions in passing that unlike derivational morphemes, inflectional morphemes may be attached higher in a tree than the X^0 node (1976: 2). He does not treat compounding, but interestingly does not comment on the decision to exclude compounding from the scope of his monograph. In other words, Aronoff's decision to discuss derivation apart from inflection and compounding does not seem to be a principled one or to have any particular theoretical significance.

Subsequent work on morphology has generally been inclusive, encompassing derivation and either compounding or inflection or both. Important dissertations such as Siegel's (1974) *Topics in English Morphology*, Allen's (1978) *Morphological Investigations*, and Lieber's (1980) *On the Organization of the Lexicon* all cover parts of the territory of morphology beyond derivation, as does subsequent influential work in word structure (Williams 1981b, Selkirk 1982, Lieber 1992), in Lexical Phonology and Morphology (Kiparsky 1982b, Halle and Mohanan 1985, Giegerich 1999), in realizational frameworks (Anderson 1992, Stump 2001), in Lexeme Morpheme Base Morphology (Beard 1995), in the onomasiological tradition (Štekauer 1998, 2005), or in the framework of lexical semantics (Lieber 2004).

Those works over the last thirty or so years that have treated derivation have tended to be focused on specific theoretical issues, for example the formal nature of rules (Aronoff 1976, Lieber 1980, 1992, Selkirk 1982, Beard 1995, Booij 2010, to name just a few), productivity (Aronoff 1976, van Marle 1985, Baayen 1989, Plag 1999, Bauer 2001), affix ordering (Fabb 1988, Hay 2000, Plag and Baayen 2009), lexicalization (Kastovsky 1982, Bauer 1983, Lipka et al. 2004), the nature of evaluative affixation (Scalise 1984, Stump 1993, Bauer 1996, 1997a, Jurafsky 1996, Grandi and Körtvelyessy forthcoming), the analysis of root-and-pattern word formation (McCarthy 1979), reduplication (Moravcsik 1978, Marantz 1982, Hurch 2005, Inkelas and Zoll 2005), and infixation (Ultan 1975, Yu 2007a). But no one seems to have taken a broad view of the subject.

1.3 A COMPREHENSIVE OVERVIEW

The chapters of this handbook thus give us a chance to ask what is distinctive about derivation. Our idea is to fill in a picture that is fragmented and currently missing key pieces. Although we have theoretical treatments of derivation, we lack a comprehensive overview that encompasses both concatenative and non-concatenative formal processes on the one

⁴ Halle (1973) draws most of his examples from derivation in English, but he briefly touches on inflection as well.

hand, and various semantic categories of derivation on the other. Further, there are surprisingly few substantial descriptive accounts of derivation in the languages of the world that allow us to make cross-linguistic comparisons; grammars of specific languages often do not have more than a few pages on derivation, and language families are almost never treated as a whole. Štekauer et al. (2012) is a step in the direction of filling in descriptive gaps, but they present isolated facts about many languages rather than focused snapshots of languages and language families. The present handbook seeks to fill this descriptive gap.

We also believe that a cross-linguistic perspective on derivation has been hampered by a view that might be too heavily Eurocentric. We give two examples. Consider the term “conversion.” This term for category change with no concomitant change in form makes sense in the context of languages like English; but it becomes increasingly problematic when we consider languages that are heavily inflected and even more so with languages that do not exhibit clear distinctions between syntactic categories (see Valera, this volume). A second example of a Eurocentric perspective might be the common notion that the formation of ideophones is not to be treated as part of derivation; current English-language textbooks on morphology (Spencer 1991, Haspelmath 2002, Booij 2007, Lieber 2010a, Aronoff and Fudeman 2011) do not even mention ideophones in the context of derivation. But the chapters in this volume on derivation in Uralic, Niger-Congo, Nilo-Saharan, and Sino-Tibetan all suggest that our view has been too narrow. In each of these families ideophones have a role to play in derivation.

Interestingly, one thing that has emerged from Štekauer et al.’s (2012) recent typological work is that it seems to be an absolute universal that languages have some sort of derivation, and this alone would justify our focus on this phenomenon. Štekauer et al. cite one language (Vietnamese) in their sample of fifty-five languages that lacks affixation, but significantly Vietnamese does not lack derivation entirely, as new lexemes in that language may be formed by various sorts of reduplication (see also Inkelas, this volume). In contrast, they cite five languages that lack compounding (Dangaléat, Diola-Fogny, Karao, Kwakw’ala, and West Greenlandic (Kalaallisut)), but that do have various formal mechanisms of derivation. The literature also suggests that some languages (Thai, Burmese, Yoruba, Vietnamese) lack inflection (Lehmann and Moravcsik 2000: 745), which would leave derivation as the only sort of morphology that all languages may be said to have.⁵ Of course this makes sense from a functional perspective: all languages need to add to their lexical stock somehow, and relying exclusively on coinage and borrowing to increase lexical stock seems implausible at best.⁶

Looking more closely at derivation, several researchers have concluded that suffixation is the most common means of derivation in the languages of the world (Hawkins and Gilligan 1988, Štekauer et al. 2012); only one affixing language in the Štekauer et al.

⁵ Greenberg (1963a) proposed the universal that “If a language has inflection, it always has derivation” (Universals Archive U506). It appears that this universal can be strengthened in light of the results we cite here: if we are correct, all languages have some sort of derivation whether or not they have inflection.

⁶ Adding to the lexical stock exclusively by borrowing may be a feature of dying languages, but is not a feature of any living language to our knowledge.

sample, Yoruba, lacks suffixation as a derivational device. Prefixation is somewhat less well-attested, although still widespread (70.91% in the Štekauer et al. sample), as are reduplication (80% in Štekauer et al., but closer to 75% in the WALS sample) and conversion (61.82% of languages in Štekauer et al.). Other forms of derivation are not nearly so widespread: Štekauer et al. say that 25.45% of the languages in their sample exhibit infixation, 21.82% circumfixation, and 23.64% stem vowel alternation (which for them includes both ablaut and root and pattern derivation). Other sorts of derivation appear in an even smaller percentage of the languages they sampled.

We therefore have some very basic knowledge of the formal, functional, and typological characteristics of derivation, but this is a bare skeleton. We intend with this Handbook to begin to fill in details in all these areas. It is our intention that the chapters gathered in this volume will be of use not only to morphologists, but also to psycholinguists, historical linguists, syntacticians, and phonologists, as well as to students and scholars in related fields that need to know about how languages add to their lexical stocks.

1.4 THE ORGANIZATION OF THE HANDBOOK

In the first part of this Handbook, we look at derivation from several perspectives. We begin with boundary issues—where to draw the line between derivation and inflection (Chapter 2) and between derivation and compounding (Chapter 3). Not surprisingly, this brings to the fore the difficulty of delineating our subject matter with perfect clarity. We next take up several “big-picture” issues including the theoretical treatment of derivation (Chapter 4), the issue of productivity and lexicalization (Chapter 5), methodologies used in obtaining data on derivation (Chapter 6), and experimental and psycholinguistic approaches to derivation (Chapter 7). Chapters 8–12 look at particular formal means of derivation (affixation, infixation, conversion, reduplication, and other non-concatenative processes). Chapter 13 looks at issues concerning allomorphy in derivation. Next, we take up derivation of nouns (Chapter 14), verbs (Chapter 15), adjectives and adverbs (Chapter 16), evaluative derivation (Chapter 17), and derivation of functional categories (Chapter 18). We also consider a number of themes that are particularly salient in the study of derivation: homophony versus polysemy in affixes (Chapter 19), paradigmatic organization in derivation (Chapter 20), and the ordering of derivational affixes (Chapter 21). Part I ends with three chapters situating derivation with respect to the wider fields of sociolinguistics, language change, and child language acquisition (Chapters 22–4).

In the second part of this volume (Chapters 25–39) we have made an attempt to fill a descriptive gap in the literature by looking at derivation across a wide range of languages. Instead of focusing on individual languages as we did in the *Oxford Handbook of Compounding*, however, we decided here to look more broadly at language families with the aim of exploring the extent of variation both within and across families. As is usually

the case in surveys of this kind, we aimed for a broad distribution of families in terms of areal and typological characteristics. Inevitably, of course, we were limited to families for which we could find willing authors. We were extraordinarily fortunate, however, in finding authors able to cover fifteen language families, ranging geographically across Europe, Eurasia, East and South Asia, Australia, the Pacific, Africa, North and South America. The reader will note that these chapters are not uniform in composition; this was inevitable, given a very wide range in the size of the language families and in the availability of data. Some chapters range broadly over many languages in the family; others give a brief overview of the family and then concentrate on one or two specific languages in the family. Chapter 34 is unique in that we could find no single author to take on all of Sino-Tibetan; this chapter is therefore divided into three sections, each covering a major branch of Sino-Tibetan. We hope that in spite of their differences in composition, these chapters nevertheless give a usefully broad overview of the range of derivation that occurs in the languages of the world.

In the last two chapters we return to broader themes. The penultimate chapter of the handbook takes an areal rather than genetic view of derivation, looking both at the mechanisms of areal spread and specific examples of areal tendencies in derivation. And in the final chapter we return to the theme of universals, assessing what the chapters of Part II of this volume can tell us about various cross-linguistic generalizations that have appeared in the literature.

We close with a word on what we have not provided in this Handbook, namely a comprehensive overview of the theoretical frameworks in which derivation has been treated. This omission was a deliberate decision on our part. On the one hand, we have already published a *Handbook of Word Formation* (2005) that covers a number of theoretical approaches to word formation. On the other, the *Oxford Handbook of Morphological Theory* (Audring and Masini forthcoming) will cover recent theoretical developments. What we hope to provide in what follows is a rich picture of how word formation works, what sorts of meanings it tends to express, how it may be studied, and how it is manifested in the languages of the world. Inevitably there will be many facets of derivation we have failed to cover adequately. Nevertheless, we hope to have provided a broad enough overview of the state-of-the-art to aid further research in the field.

CHAPTER 2

DELINEATING DERIVATION AND INFLECTION

PIUS TEN HACKEN

THE distinction between derivation and inflection is one of the traditional problems of linguistic morphology. Although the concepts are intuitively clear, the boundary between them is elusive when borderline cases are considered. Here, I will start by presenting the intuitive core of the opposition (Section 2.1). Then some general considerations from the theory of terminology are discussed, which determine the framework of discussion (Section 2.2). Within this framework, there are two main positions that have been taken, one that there is a categorical opposition, the other that any attempt to define the two categorically is futile (Section 2.3). Against this background, I will then discuss some criteria that have been proposed (Section 2.4) and some problem cases for the classification (Section 2.5).

2.1 THE CORE OPPOSITION

Both inflection and derivation are concerned with morphologically related forms. A clear example of inflection is the set of case and number forms of Polish *kobieta* (“woman”) in (1).

(1)	<i>Singular</i>	<i>Plural</i>
<i>Nominative</i>	kobieta	kobiety
<i>Genitive</i>	kobiety	kobiet
<i>Dative</i>	kobietcie	kobietom
<i>Accusative</i>	kobietę	kobiety
<i>Instrumental</i>	kobietą	kobietami
<i>Locative</i>	kobietcie	kobietach
<i>Vocative</i>	kobieto	kobieto

In (1), there are ten different forms occupying fourteen case-number slots. Most nouns in Polish have the same set of fourteen slots illustrated in (1). Together, the forms in (1) are called the *paradigm* of *kobieta*. The paradigm together with the citation form is called the *lexeme*, e.g. by Matthews (1974: 21–2).

A clear example of derivation is the English pair in (2).

- (2) a. read
b. readable

The pair in (2) has a number of properties that make it a typical example of derivation. Whereas *read* in (2a) is a verb, *readable* in (2b) is an adjective. There are various ways in which this pair differs from the paradigm in (1). Perhaps the most significant one is that the pair in (2) is not a paradigm of a single lexeme, but represents the incidental formation of a new lexeme.

The contrast between (1) and (2) can be taken to be prototypical for the distinction between inflection and derivation. In this case, many properties can be used to classify (1) as inflection and (2) as derivation. However, there are many instances in which the contrast is less obvious. The cluster of properties that distinguish (1) and (2) tends to disintegrate when we consider borderline cases. The discussion of whether and how to delineate inflection and derivation concentrates on such cases, using them either as an illustration of where the boundary should be drawn or as an argument against drawing a categorical boundary and see the contrast as a continuum instead.

2.2 TERMINOLOGICAL CONSIDERATIONS

The problem of determining the precise extent of the categories of inflection and derivation is an instantiation of a general terminological problem. Natural concepts are prototypes. Here the expression *natural concept* refers to a concept as it emerges in a speaker's mind on the basis of usage and exemplification. Terminological concepts, that is concepts with what Bessé (1997) calls a "terminological definition," are unnatural in their categorical delimitation.

Labov (1973) demonstrated that a natural concept such as *cup* has fuzzy boundaries by studying the categorization judgments for objects ranging from clear cups to clear bowls. As Jackendoff (1983: 86) observes, such judgments must be based on the application of rules, because we do not learn the category of each object separately. These rules are unconscious and they constitute an important part of the meaning of the relevant concept. In the case of *cup* and *bowl*, Labov (1973) identifies two types of condition that are responsible for the gradual transition between them. First there are scalar conditions such as the height–width relation. Secondly, there are what Jackendoff (1983: 137) calls *preference rules*, such as the property of having a handle. Preference rules are neither necessary nor sufficient, but they interact with scalar conditions so that, for instance, an

object that because of its height–width relation is rather a bowl may be judged rather a cup when it is given a handle.

The idea that natural concepts are prototypes is elaborated by Rosch (1978) for general language, but Temmerman (2000) argues that it also applies to terminology. Her examples are from the domain of the life sciences, but the insights can be applied to the concepts of inflection and derivation as well. Discussing legal terminology, ten Hacken (2010a) argues that the questions this situation raises are whether or not it is worth formulating a terminological definition in the sense of Bessé (1997) and if yes, how to do so. These questions are equally relevant to the linguistic concepts of inflection and derivation.

As explained by ten Hacken (2010a, b), formulating a terminological definition is equivalent to creating an abstract concept. In the case of legal concepts, such definitions are necessary for the enforcement of laws. Without a proper definition of *parking* in traffic law, constraints on parking cannot be enforced. In scientific terminology, discussed by ten Hacken (2010b), the question is whether the concept contributes to the explanatory power of the theory it is used in. A linguistic example illustrating the relevance of this question is the notion of *subject* in relation to German (3).

- (3) Mir ist kalt.
 Me_{DAT} is cold
 i.e. ‘I am cold’

It is not immediately obvious whether *mir* in (3) is a subject. The question is whether this is a problem. In Lexical-Functional Grammar (Bresnan 2001), with its separate level of *f*-structure in which grammatical functions such as *subject* are primitives, it is essential to define *subject* exactly. We need to know whether *mir* in (3) is a subject or not in order to produce a correct *f*-structure. In Head-Driven Phrase Structure Grammar, at least in the version presented in the first eight chapters of Pollard and Sag (1994), subjects are not formally distinguished from other complements, so that there is no need to define *subject* as an abstract object. In the representation of (3), *mir* is on the subcategorization list, but it need not be specified whether it is the subject or not. This does not mean, of course, that Pollard and Sag (1994) claim that there are no differences at all between subjects and other complements. The contrast between Bresnan (2001) and Pollard and Sag (1994) in this matter only concerns the theoretical significance of these differences. As we will see in Section 2.3, the same type of discussion can be found in the context of inflection and derivation.

If we decide to set up a terminological concept, the next question is how we select the conditions in the definition. For scientific terminology such as *inflection* and *derivation*, terms are part of a network of abstract concepts imposed on reality. Links in the network are references to a term in the definition of another term. A good definition of a concept is one that contributes to making the network of concepts a good basis for an explanatory theory.

Given the aim of increasing depth and scope of explanation by scientific theories, it is inherent in the history of scientific concepts that they have to adapt to extensions in the empirical and theoretical basis. This can be illustrated by the history of the term *planet* in astronomy (cf. Schilling 2007). In 17th- and 18th-century astronomy, it was sufficient to define a planet as a body in orbit around the Sun that does not emit but only reflects light. Equivalently, at least until Uranus was discovered in 1781, the six planets could be simply listed. The discovery of increasing numbers of small planets in the 19th century led to the creation of a new concept *asteroid*, distinct from *planet*. It is important to see the relation between the empirical basis, the theoretical basis, and the terms here. The extension of the empirical basis was in principle easily accommodated by means of the existing terms, but it triggered a theoretical need to distinguish a new concept. Similarly, Schilling describes how the discovery of Pluto in 1930 was at first accommodated by classifying it as a planet, but when further discoveries were made this decision was revised, leading in 2006 to the International Astronomic Union (IAU) definition of *planet* in terms of necessary and sufficient conditions.

In general, we cannot assume that the definition of a term will persist over time. It is natural that changes in theory and knowledge lead to different, more advanced definitions. In the field of terminology, the need for regular updates of definitions of terms is recognized (cf. Arntz et al. 2009: 69). In the case of terms such as *inflection* and *derivation*, we are dealing with concepts that have clearly distinct prototypes, as (1) and (2) illustrate, but are at the same time placed in a continuum of more or less typical cases. When delineating the concepts, the best we can do is to draw the borderline so that it runs through a (relatively) sparsely populated area of the continuum and uses theoretically significant properties. However, extensions of the empirical basis can increase the number of borderline cases and theoretical developments can shift the emphasis away from properties once thought to be significant.

2.3 TWO APPROACHES

Approaches to the distinction between inflection and derivation can be divided into two types, each with a rather long tradition. I will call them here the *categorizing tradition* and the *skeptical tradition*. In the categorizing tradition, the position is that inflection and derivation should be treated as different categories and the boundary between them should therefore be clear. In the skeptical tradition, we find two patterns of reasoning that converge on the same result. In one, it is argued that a clear boundary between inflection and derivation cannot always be achieved, so that we should formulate our theories in such a way that it is not necessary. In the other, it is argued that the best theory does not depend on the distinction between inflection and derivation, so that there is no reason to try to define this distinction precisely. The two reasonings are often used to reinforce each other, both leading to abandoning the search for precise criteria to delineate inflection and derivation.

In traditional grammars, inflection is a central topic of the grammar, whereas derivation is not included and is taken to be covered by the dictionary. We find this in Hoffmann's (1777) grammar of Latin, Girauld Duvivier's (1822) grammar of French, but also in Bornemann and Risch's (1978) grammar of Ancient Greek. The dominance of inflection in this type of grammar is illustrated by the space devoted to different sections. Bornemann and Risch (1978) devote 25 pages to phonology, 136 pages to declension and conjugation, and 144 pages to what is called "Syntax," but the first 107 of these pages are about the choice of the correct inflected form of words in a particular context. Older grammars tend to discuss orthography instead of phonology, but the pattern is otherwise very similar. These grammars often include appendices. Thus, Hoffmann (1777) has an appendix on the Roman calendar. Bornemann and Risch (1978) include a 15-page appendix on Greek word formation, alongside one on the Homeric language and one on Greek meter. Significantly, Bornemann and Risch (1978) do not discuss the distinction between word formation and inflection at all, apparently taking it as given.

The approach to delineating inflection and derivation in traditional grammars can be compared to the approach to the concept of *planet* in 17th- and 18th-century astronomy. Inflectional categories, like planets, were defined by means of a list or some general descriptive properties and the two ways of defining them were taken to be equivalent. The listing approach requires that either the categories of one language (e.g. Latin) are taken to be universal, or that each language is considered as a separate universe. A grammar such as Guasch (1983) for Guaraní is an interesting mix between the two. Thus, Guasch (1983: 51–3) first states and exemplifies that nouns are not inflected for number and gender, before treating their inflection for tense. This approach can be explained (and justified) by the use of traditional grammars in language teaching.

The skeptical tradition emerged as a reaction against the position adopted in traditional grammars, that is that the boundary between inflection and derivation is obvious. Bloomfield (1933: 223–4) starts his overview of criteria by which inflection has been distinguished from derivation with the remark that "[t]his distinction cannot always be carried out." What Bloomfield means is that in some languages and for some morphological constructions, it is not possible to determine whether they are inflectional or derivational. A stronger formulation of this position is the one by Bloch and Trager (1942: 54), given in (4).

- (4) For some languages, it is useful to divide the morphological constructions of complex words into two kinds according to the grammatical function of the resulting form: DERIVATIONAL and INFLECTIONAL.

Whereas Bloomfield presents the question of whether inflection and derivation can or should be distinguished as a matter of debate, in (4) the scope of the distinction is restricted to "some languages." In interpreting these statements, it is important to keep in mind the nature of the text they appear in. Bloomfield's book is an overview of

linguistic analysis “intended for the general reader and for the student who is entering upon linguistic work” (1933: vii). We might call the book a textbook. This explains the implication of a debate. Bloch and Trager write in their introduction that their aim is “to present in brief summary the techniques of analysis which are necessary for learning a foreign language by the method of working with native speakers and arriving inductively at the grammatical system of their language” (1942: 4). It does not give a full overview of the state of the art in linguistics, but is intended as a guide for language learners. Therefore, (4) does not imply any debate, but just describes the usefulness of the distinction in “some languages.”

In early generative grammar, there was no obvious place for morphology. In Chomsky’s (1957) model, syntax is governed by rewrite rules and transformations that operate on morphemes. The phonetic realization of these morphemes is attributed to interpretation rules operating on Surface Structure, whereas semantic interpretation rules operate on Deep Structure to produce the representation of meaning. In such a model, there is no basis for any distinction between inflection and derivation. Nominalization transformations such as proposed by Lees (1960) are formally of the same type as Chomsky’s (1957: 39) transformation that produces the past tense of verbs. Inflection and derivation are at most pre-theoretical, descriptive terms in such a theory.

The analysis of nominalization was a crucial battleground in what Paul Postal called the “linguistic wars” (Newmeyer 1986: 117). In Generative Semantics, nominalization was accounted for by means of transformations. Nominalization was also taken in a much broader sense. Thus, Levi (1978: 168) classifies both *city planner* and *car thief* as agent nominalizations. The reason is that she assumes that *thief* includes the predicate also found in *steal* in its Deep Structure. In Generative Semantics, we can therefore observe a continuation of the early generative position that morphology is dealt with by means of syntactic rules, which does not give any reason to distinguish inflection and derivation.

The opponents of Generative Semantics made use of the lexicon, introduced by Chomsky (1965: 84–8) as a part of the base component, alongside the rewrite rules. The base component generates Deep Structure and the lexicon contains a specification of “all properties that are essentially idiosyncratic” (1965: 87). Chomsky (1970) argues for the “lexicalist hypothesis,” which implies “that derived nominals will correspond to base structures rather than transforms” (1970: 193), that is, they are in the lexicon rather than the result of syntactic rule application.

Obviously, Chomsky (1970) uses nominalization as an example, but it is not unequivocally clear how far this example should be extended. Whereas it is straightforward to extend the scope of the treatment proposed for nominalization to other types of derivation, the question of whether it should be extended to include inflection remains open. Scalise (1984: 101) uses the terms *Strong Lexicalist Hypothesis (SLH)* and *Weak Lexicalist Hypothesis (WLH)* to distinguish these options. The choice between them has implications for the distinction between inflection and derivation. In the WLH, only derivation is in the lexicon, whereas inflection is covered in syntax and/or phonology. Therefore, inflection and derivation *must* be distinguished in a categorical way. In the SLH, both

inflection and derivation are covered in the lexicon. They may be distinguished, but the status of the distinction is not determined by the grammatical framework.

Two foundational texts elaborating the WLH are Aronoff (1976) and Anderson (1992). Both start from the assumption that derivation creates new lexemes, whereas inflection generates the paradigm of word forms of a lexeme. Both reject the morpheme as the basic unit of morphology. Aronoff (1976: 115) makes the claim in (5).

- (5) [M]orphology is word-based: new words are formed from already existing ones, rather than being mere concatenations of morphemes.

In interpreting (5) it should be taken into account that Aronoff uses *word* where Matthews (1974) would use *lexeme* (cf. Aronoff, 1976: xi). Anderson (1992) subscribes to the claim in (5) and develops a system of *A-Morphous Morphology*, that is one where morphemes do not play any role. In these systems, the lexeme (or word, in Aronoff's terminology) is the anglepoint between inflection and word formation, so that it is crucial to distinguish inflection and derivation precisely.

One approach that continues the skeptical tradition is Distributed Morphology (DM). Its foundational text is Halle and Marantz (1993). Harley and Noyer (2003) present a general overview and Harley (2009: 131–3) gives an update. The general idea of DM is that syntactic structure reaches all the way down to the level of the morpheme. As Harley and Noyer (2003: 474) state, the distinction between inflection and derivation “has no explicit status in DM,” but there is a distinction between functional and lexical morphemes (*f-morphemes* and *l-morphemes*) which expresses some of the difference between prototypical cases such as (1) and (2). In syntactic structure, all morphemes are feature bundles, but *f-morphemes* and *l-morphemes* operate differently in mapping them to phonological representations. In *f-morphemes*, all vocabulary items are in competition and rules are devised to select the right one, whereas for *l-morphemes*, the choice is between different lexical items with different encyclopedic content.

It is interesting to compare this approach to Lieber's (2004). Lieber also assumes that morphology is a theory of morphemes, but she quite explicitly distinguishes inflection and derivation. Lieber's morphemes are composed of a skeleton and a body (2004: 9), where the skeleton contains the more formalized features and the body the encyclopedic information. Lieber distinguishes inflectional and derivational affixes on the basis of the contribution they make to the meaning of the base they attach to (2004: 151). Derivational affixes are morphemes that have an argument in the skeleton, so that they change the referential meaning of the base, whereas inflectional affixes lack such an argument. It should be noted, of course, that this is intended as a way of representing the difference, not of making the distinction. As opposed to the situation in word-based and *a-morphous* morphology, the distinction is not itself crucial for Lieber's framework.

In presenting her framework for lexical semantics, Lieber refers to Jackendoff's Lexical Conceptual Structures as the basis for her formalism of the skeleton. However, in elaborating his Parallel Architecture, Jackendoff (2002: 152–62) argues for the complete abolition of the traditional distinction between inflection and derivation. As an example of what is

usually treated as inflection, he discusses the English past tense (2002: 160–2). The past tense ending *-ed* is for Jackendoff a lexical entry of its own, specifying in its phonological information that it is an affix, in its syntactic information that it attaches to a verb and makes it tensed, and in its conceptual information that it marks the past. Strong verbs such as *eat* have separate lexical entries for the stem and for the past tense. He mentions “massively affixing languages like Turkish” as an argument in favor of this approach (2002: 156).

Jackendoff’s theory not only rejects (5), but even abandons the notion of *lexeme*. This raises the question of how to express the regularity of the pattern in (1). Booij (2010) presents Construction Morphology as a morphological theory within Jackendoff’s general architecture. He proposes to represent inflectional paradigms as correspondence relations between constructional schemas (2010: 255–7). These relations can be encoded as redundancy rules (cf. Jackendoff 1975) so that the pattern in (1) is stored as one of the typical ways of generating the nominal paradigm in Polish. Redundancy rules cover emergent patterns and facilitate lexical storage and retrieval, but they are not crucial for generating correct expressions.

Instead of lexeme formation and lexeme realization, Jackendoff only distinguishes productive and semiproductive affixes. The latter cover all cases where limitations on the regular formation of expressions cannot be predicted on the basis of conditions that can be encoded in the relevant lexical entry. Jackendoff (2010: 34) identifies semiproductivity as “one of the central issues of linguistic theory for the coming years.” Semiproductivity is in principle independent of the distinction between inflection and derivation, as noted by Jackendoff (2002: 155).

In conclusion, the status of the distinction between inflection and derivation is a consequence of theoretical assumptions. There are two main approaches in this respect. One continues the traditional distinction made in school grammars and highlights the importance of lexemes and paradigms, but aims to give it a stronger terminological foundation. The other is skeptical about the possibility of doing so reliably. It tends to highlight the difficulties of classifying borderline cases. However, even if they do not require a precise distinction, most frameworks at least provide for a way to encode the general prototypes underlying the differences illustrated in (1) and (2).

2.4 CRITERIA FOR THE DISTINCTION

Given the terminological status of *inflection* and *derivation*, we can expect as the main sources for the discussion of the distinction between them texts of three types. First, sections of textbooks introducing students to the field of morphology. Secondly, sections of handbooks giving an overview of the field. Thirdly, argumentative articles or sections of monographs presenting or discussing frameworks in which the distinction plays a crucial role. The first two of these are generally the most prolific in the use of terms (cf. Pearson, 1998). They reflect communication types in which terminology is typically introduced and explained. The last one is a sign of the controversial nature of the distinction and is an important source of defining criteria.

Modern textbooks, for example Aronoff and Fudeman (2005) and Fábregas and Scalise (2012), typically devote only a few pages to the distinction. The textbook nature of the former is reflected in the division of the material between a short section introducing the intuitive notions with some examples as part of the introduction of the notion of *lexeme* (2005: 44–6) and an overview of the main distinguishing criteria as part of the discussion of inflection (2005: 160–3). Fábregas and Scalise (2012: 104–8) only give some examples suggesting that the distinction is problematic and the two concepts should be seen as prototypes. Earlier textbooks, for example Scalise (1984: 102–15) and Bauer (1988b: 73–87), present much more substantial overviews of the criteria used. Scalise and Bauer take explicit but opposing positions as to the status of the distinction. Scalise (1984: 103) announces at the outset that “we will argue in favor of the division,” whereas Bauer (1988b: 85) concludes that “[n]one of the criteria has appeared satisfactory.”

Handbooks are less pedagogically oriented, but give a more systematic overview of the field. The division of morphology into topics influences how the distinction between inflection and derivation is treated. Spencer and Zwicky (1998) include separate chapters on inflection and derivation, each of which addresses the distinction between them. Both Stump (1998: 14–19) and Beard (1998: 44–6) list criteria that have been used, give examples of problems for the classification, and address the issue of how the distinction should be interpreted in the light of these problems. Booij et al. (2000) devote chapters to the borderlines between the phenomena. Booij’s (2000) discussion of inflection and derivation follows the same pattern as Stump’s (1998) and Beard’s (1998), but goes into more detail. Müller et al. (forthcoming) concentrate only on word formation, so that the question is one of delimiting the scope of the volume. Compared to the other discussions, Štekauer’s (forthcoming a) stands out because it starts with an overview of the reasons why the boundary is hard to determine before giving an overview of criteria. All of them have a rather skeptical view of the feasibility of the distinction.

Stump (1998: 14) observes that the criteria he presents are logically independent and “one wouldn’t necessarily expect each of the five criteria to divide morphological phenomena into the same two groups.” This sums up very well the terminological problem of turning prototypes into precise concepts. The strength of the prototype is the result of converging criteria, but when these criteria are used in a definition, the differences between the sets of phenomena they identify are highlighted. As noted by Bessé (1997), in formulating a terminological definition, choices have to be made.

The final category of sources includes those in which a technical solution to the practical problem of distinguishing the two categories is presented. A well-known example is Anderson (1992), whose theory takes inflection to be in a different part of the grammar from derivation. Another example is ten Hacken (1994), who approaches the problem from the perspective of Word Manager (cf. ten Hacken, 2009), a system for electronic morphological dictionaries in which lexemes are the basic units of description. In both cases, a critical discussion of the criteria that have been used is followed by a solution. Anderson (1992: 82–5) summarizes what he calls the “substance of the notion of inflection.” Ten Hacken formulates independent terminological definitions for inflection (1994: 298) and derivation (1994: 303) on the basis of his discussion.

I will now turn to a number of commonly used criteria. For reasons of space, I cannot present all criteria referred to in the sources mentioned above. Given the large overlap between discussions, I will only give individual references where there is a reason to single out one approach from among the others.

A frequently used criterion is based on the relative order of affixes, formulated by Greenberg's (1963b: 93) as (6), number 28 of his universals.

- (6) If both the derivation and inflection follow the root, or they both precede the root, the derivation is always between the root and the inflection.

In (6), "derivation" and "inflection" refer to the relevant affixes. As a generalization about word forms that include both types of affixes, (6) is quite strong, but not without apparent exceptions. An example of a problem case is the formation of adverbs in French, illustrated in (7).

- (7) a. lent ('slow') base form and masculine singular
 b. lente ('slow') feminine singular
 c. lentement ('slowly')

The adverb (7c) seems to be derived from the feminine form (7b). Historically, such an analysis is indeed correct because Late Latin *mentem* ('character, manner') is a feminine noun. In order to reconcile the data in (7) with the generalization in (6), we would have to claim that *lente* in (7c) is not an inflected form of (7a), but the base form or a stem variant, or that French adverb formation is inflectional.

Apart from empirical problems, ten Hacken (1994: 155–6) also notes a technical problem with (6). If we have a word form with two affixes, for example Base-Affix₁-Affix₂, (6) can only be applied to determine the category of an affix if we already know that Affix₁ is inflectional or that Affix₂ is derivational. The inflectional status of Affix₁ or the derivational status of Affix₂ must be established on the basis of other criteria. Therefore, (6) can at most be an auxiliary criterion.

Another frequently used criterion is based on the syntactic category of the base and the output. Scalise (1984: 103) formulates it as (8).

- (8) I[nflection] R[ule]s never change the syntactic category of a word, while
 D[erivation] R[ule]s may change it.

The contrast between (1) and (2) provides a good example of (8). Obviously, (8) depends on an independent definition of *syntactic category*. In the context of (7), it is relevant that it has been argued that adverbs such as *slowly* are inflected adjectives, for example by Hockett (1958) and by Larson (1987). Another problem is the classification of participles (cf. Section 2.5). Technically, it is not a problem if one concept is dependent on another. The terminology of a particular field can often be seen as a network of terms

related to and referring to each other. More problematic is that (8) is formulated as only a sufficient condition for inflection. If adjectives and adverbs are separate syntactic categories, (8) tells us that the formation of (7c) is derivation. If they are not separate categories, (8) does not tell us anything. We have to admit that derivation sometimes does not change the syntactic category, for example in the prefixation in (9).

- (9) a. clear
b. unclear

It is obvious that *un-* changes the meaning of the base it attaches to in a way that is very similar to the typically derivational contrast in (2) and different from the typically inflectional contrasts in (1). However, the syntactic distribution of (9a and b) is so similar that it is almost impossible to argue that they belong to different syntactic categories. Although Scalise (1984: 103) suggests that “[t]here are reasons...for believing that a DR always changes the syntactic category of its base,” he only gives examples such as *avvocato* (‘lawyer’) and *avvocatura* (‘lawyership’), where countability and abstractness features change. Scalise (1984: 109–10) also gives inflection class, subcategorization, and selectional features, $\pm$ animate and $\pm$ common as relevant features. However, *un-* in (9b) does not change any of these. We can only observe that the meaning it contributes is rather different from the case affixes in (1). Therefore, the existence of many cases such as (9) reduces the value of the criterion in (8) for delineating inflection and derivation.

A third widely used criterion is based on productivity. Aronoff and Fudeman (2005: 161) formulate it as in (10).

- (10) [I]nflectional morphology tends to be more productive than derivational morphology.

As formulated, (10) raises two problems, both of a by now familiar nature. First, the hedge “tends to” and the degree “more” make (10) a characterization of the prototypes rather than a criterion to be used in a terminological definition. Second, the use of productivity makes (10) dependent on a definition of this concept. *Productivity* has been used in different senses and for our purposes Corbin’s (1987) analysis into three concepts is useful. The underlying notion of *disponibilité* (‘availability’) does not distinguish inflection and derivation, because both consist of a large body of available affixes or processes. The derived notions of *rentabilité* (‘profitability’) and *régularité* (‘regularity’) are more interesting here. *Rentabilité* is a gradual property and is realized to the highest degree when it can be reliably predicted that the output of the process exists. The idea of *régularité* is that the resulting word (or word form) has a predictable form and meaning.

English nominal plural /z/ is a good example of a highly productive process on both counts. It applies to almost all nouns unless there are semantic reasons for not having a plural. Only very few nouns form their plural in other ways. Moreover, the form and meaning are in almost all cases entirely predictable. There are three phonological realizations of /z/, but the choice among them is entirely determined by the last phoneme of

the base. Apart from a few lexicalized plurals with special meanings, the meaning is the combination of “plural” with the meaning of the base noun. This makes it a prototypical case of inflection. However, irregular inflection, for example *-en* as an English nominal plural, scores low on both counts. There are very few cases where it applies and in the case of *children*, it triggers further, unpredictable phonological changes. This arguably makes *-en* more typical of derivation.

When we apply (10) to (1) and (2), we encounter a different type of problem. For most case-number combinations, Polish has different possible endings and the choice among them can only be predicted in part by phonological properties, gender, and animacy of the base noun. A well-known problem case is the formation of the genitive singular of masculine inanimate nouns, as illustrated in (11).

- (11) a. *ser* *sera* (‘cheese,’ nom./gen. sg.)
 b. *deser* *deseru* (‘dessert,’ nom./gen. sg.)

There is no general rule saying when *-a* or *-u* is to be used. Polish grammars, for example Bielec (1998: 109–10) and Orzechowska (1999: 306), give semantically based generalizations, but they are not absolute rules. Conversely, the pair in (2) is an example of a highly productive affixation process in English. Almost all transitive verbs can have an adjective in *-able* with the meaning “which can be V-ed.” On the basis of such considerations, Bauer (1988b: 79–80) argues that “derivation is more productive than is generally thought,” whereas “[i]nflection is less productive than is frequently believed.”

A possible way out in view of data such as (11) is to assume that the unit for which we determine whether it is productive or not is not the affix, but the feature combination. Every Polish noun has word forms for each of the slots illustrated in (1), except if there are obvious semantic reasons for not having a plural. This would also solve the problem of classifying irregular plurals in English as inflection. This is the basis of Matthews’ (1974) *Word and Paradigm* model. The idea is that inflection has paradigms but derivation does not.

There are two types of problems with this idea. The first is the existence of so-called *defective paradigms*. Thus, for the present indicative of the French verb *clore* (‘close’), Grevisse (1980: 810–11) gives only the forms in (12):

(12)		<i>Singular</i>	<i>Plural</i>
	<i>First person</i>	<i>je clos</i>	—
	<i>Second person</i>	<i>tu clos</i>	—
	<i>Third person</i>	<i>il clôt</i>	<i>ils closent</i>

Despite paradigmatic pressure, there are no forms for the first or second person plural. Yet, the forms in (12) are prototypically inflectional. A much more serious problem with paradigms as a criterion to distinguish inflection and derivation, however, is of a general terminological nature. In order to use *paradigm* in the definition of inflection, we should have a definition of *paradigm* that is independent of *inflection*. In Latin grammars, verbs are neatly organized in conjugation classes with forms in each slot representing a feature

combination. However, when we only have as a basis the set of word forms, for instance for an as yet undescribed language, and have to determine which features constitute the structure of the paradigm, it is by no means straightforward what should be included in the paradigm. As Anderson (1992: 79–80) notes, it is difficult to escape circularity of definitions here.

Whereas all of the criteria discussed so far may serve to illustrate the nature of the prototypes of inflection and derivation, they have drawbacks when used as the basis of a terminological definition. In a context in which a definition of that type is required, Anderson (1982: 587) proposes (13) as the starting point:

- (13) Inflectional morphology is what is relevant to the syntax.

It is important to understand the status of (13). Bauer (1988b: 84–5) claims that “it is not sufficient as it stands to define the precise area it wishes to capture,” noting, for instance, that different syntactic theories lead to different sets of properties being relevant. In the original context of Anderson (1982, 1992), however, (13) is only the slogan used as a headline for a more precise claim supported by an elaborate theory that specifies what is relevant to syntax and why. For instance, the change of category in (2), though undoubtedly “relevant to the syntax” in a general sense, is not in the scope of (13). The only valid point Bauer can be said to make (or at least imply) here is that a terminological definition of inflection is theory-specific. This is true for scientific terminology in general and can therefore not be used as an argument against any specific definition.

A central element of Anderson’s system is the notion of agreement. The contrast in (14) can serve as a starting point.

- (14) a. One delegate from each country attends the meeting.
b. Two delegates from each country attend the meeting.

The different forms of the verb *attend* in (14a) and (14b) do not indicate properties of the verb, but only properties of its subject. Therefore, the form of the verb is not a lexical choice, but it depends on agreement.

Anderson (1992: 82–3) distinguishes four types of relevant properties. They are illustrated in (15).

- (15) a. Agnieszka cieszy się nową sukienką.
Agnieszka_{NOM} is.happy.about REFL new_{FEM-INSTR-SG} dress_{INSTR-SG}
i.e. ‘Agnieszka is happy about her new dress’
b. Ankara ve İzmir gideceğim.
Ankara and Izmir_{DAT} I.go
i.e. ‘I go to Ankara and Izmir’

Anderson’s first type of inflectional property he calls configurational. In Polish (15a) we find this when the verb *cieszyć się* (‘be happy about’) governs the instrumental case

of *sukienka* ('dress'). The case, number, and gender of *nowy* ('new') are determined by agreement, Anderson's second type. The feminine gender of *sukienka* is an inherent feature, Anderson's fourth type. His third type is phrasal properties. An example is the dative ending *-e* in the Turkish example (15b). This ending has scope over the entire coordinated NP *Ankara ve İzmir*, so that the first of these does not get any case ending.

It should be kept in mind throughout that the classification as inflection or derivation pertains to features, not to individual occurrences. The fact that the singular number of *meeting* in (14) or the feminine gender of *Agnieszka* in (15a) does not trigger agreement in these sentences is not relevant. The point is that there are contexts in which these features trigger agreement, for example for *delegate* in (14) and for *sukienka* in (15a).

Whereas in distinguishing inflection and derivation Anderson (1992) concentrates on identifying properties of inflection, ten Hacken (1994) proposes independent definitions of inflection and derivation. The definition of inflection (1994: 298) is (16).

- (16) An inflection process is a process realizing a feature or combination of features F on a word W, such that:
- The value of F is determined by agreement with another word or with a functional category.
 - If the two elements in agreement are in X and Y, either X and Y are in the same maximal s-projection, or the maximal s-projection of Y is the complement or the specifier of X.

It is noteworthy that (16) is formulated as a terminological definition in Bessé's (1997) sense. Compared to Anderson (1992), it relies more heavily on agreement. The technical formulation is meant to unify Anderson's configurational and agreement properties into one class. The final clause is meant to distinguish inflection from certain types of clitics. The term *maximal s-projection* refers to a domain of agreement that prevents, for instance, French pronominal clitics from being analyzed as inflectional markings. As noted above, ten Hacken's (1994) definitions are intended to be used in the context of Word Manager. This framework treats clitics in a different way to inflection because it takes the lexeme in the sense of Matthews (1974) as the basic unit of description. As a consequence, Anderson's (1992) category of phrasal properties is not recognized as inflectional. His category of inherent properties are not included in inflection because they are not features that need to be realized. Ten Hacken's (1994: 303) definition of derivation is (17).

- (17) A derivation process is the application of a functor element F to a word or phrase W in the lexicon, such that:
- The relation between W and F(W) can be expressed in terms of modification of the argument structure and/or the syntactic category of W;
 - For any W', if F can apply to W', the relation between W' and F(W') is the same as the relation between W and F(W);
 - Neither F nor W can play an independent role in syntax, but only F(W) can do so.

The idea that derivation is defined independently is remarkable, because in general the discussion of the way to delineate it from inflection concentrates on properties of inflection. Inspired also by Anderson (1992), (17) takes a process-based view of derivation, but whereas inflection realizes features, derivation brings about semantic and/or syntactic changes to the base. The second clause states that the derivational operation must have the same effects on different bases. The base can be a word or a phrase and, according to the final clause, it is not itself available for pronominal reference or other syntactic operations. This can be seen as the effect of the output ending up in the lexicon. The type of operation is restricted by the condition in the first bullet point. As it stands, it is not obvious how prefixation as in (9) is included in the scope of derivation, but there are various ways the clause could be amended to remedy this.

Anderson's (1992) delimitation of the domains of syntax and the lexicon and ten Hacken's (1994) terminological definitions of inflection and derivation illustrate how the categorical approach has been pursued. The perceived success of such approaches depends on the tolerance to the use of theory-internal concepts and to individual classification decisions that do not converge with traditional classifications.

2.5 SOME BORDERLINE CASES

Among the phenomena that have been treated as derivation by some and as inflection by others are adverbs, participles, and diminutives. The first two of these put into question the notion of lexeme as used in traditional grammars of Latin and Greek.

In the case of adverbs and participles, the issue is the set of syntactic categories. As noted in the discussion of (8), change of syntactic category is one of the most commonly adopted criteria for delineating inflection and derivation. The status of adverbs was mentioned in the discussion of (7) above. Whereas classical grammarians consider them a separate category, some modern theories take them to be inflected forms of adjectives. In the case of participles, classical grammarians such as Dionysios Thrax treat them as a separate category (cf. Robins 1979: 33–4), but from the 18th century onwards traditional grammars of Greek and Latin include them in the verbal paradigm. A special case is found in Celtic languages, where so-called *verbal nouns* are by far the most frequent form of verbs. In her detailed analysis of verbal nouns in Irish, Bloch-Trojnar (2006) argues that two of their four main uses are inflectional and the other two derivational. This is comparable to analyzing past participles such as (18a) as inflectional, but attributive passive participles such as (18b) as derivational.

- (18) a Boris has left his luggage at the railway station.
b. The problem of left luggage was discussed at the meeting.

How attractive a split analysis of the participle is, depends on the theoretical framework adopted. Bloch-Trojnar (2006) adopts Beard's (1995) Separation Hypothesis, which

radically separates the formation of a word form from its syntactic and semantic interpretation. In a framework in which a stronger correspondence between form and meaning is assumed, it is problematic to consider *left* as both inflectional and derivational when its irregular formation is the same in both cases.

Diminutives and augmentatives are addressed in more detail in another chapter of this volume. Here they are mainly interesting for the cross-linguistic differences in status. Whereas in Indo-European languages they are derivational, Anderson (1992: 80–1) notes that in Fula they behave inflectionally. Not only are they fully regular, Arnott (1970: 92) also gives examples of agreement such as (19).

- (19) a. loo-nde 6alee-re (<a> black storage-pot')
 b. loo-d'e 6alee-je ('black storage-pots')
 c. loo-ŋgel 6alee-yel (<a> little black storage-pot')
 d. loo-kon 6alee-hon ('little black storage-pots')

In (19), we see that the noun and adjective agree not only in number, but also in the feature diminutive. It is not the color referred to by the adjective, but the object referred to by the noun that is diminutivized. This is the same as the agreement of *nowq* in number, gender, and case in (15a). The agreement in (19) provides a strong argument for considering diminutives in Fula inflectional, whereas they are derivational in, for instance, Russian and Italian. Cross-linguistic variation of this type can occur whenever we have a feature that can be construed as meaningful, but also as a purely grammatical feature. Another feature which displays such variation is number, which is inflectional in Indo-European languages, but not, for instance, in Chinese (cf. Wiedenhof 2004: 217).

Phenomena at the borderline between inflection and derivation are often invoked as an argument that inflection and derivation should be seen as endpoints of a continuum. If we want to preserve inflection and derivation as concepts about which theoretical claims can be made, we need to select criteria as part of a terminological definition. Such a definition will then determine whether they are inflection or derivation.

CHAPTER 3

DELINEATING DERIVATION AND COMPOUNDING

SUSAN OLSEN

3.1 INTRODUCTION

THE *Handbook of Derivational Morphology* aims to provide insight into the derivational means of vocabulary extension found in natural language. Apart from overt affixation (i.e. suffixation, prefixation, circumfixation, infixation, transfixation, etc.), these means include conversion, back-formation, analogy, truncation, blending, and reduplication. Derivational morphology together with compounding constitutes the field of word formation which studies the creation of new lexemes. Inflectional morphology examines the (declensional or conjugational) variation in form of existing lexemes and is the topic of Chapter 2 in this handbook. This chapter concentrates on the delineation of the two major categories of word-formation, derivation and compounding, in order to provide a clearer vision of the type of phenomena that fall under consideration as products of derivational morphology.

Compounding, simply spoken, is a combinatorial word-formation process that creates complex words by combining lexemes (roots or stems). Its products, that is, compounds, are comprised of two or more lexemes at the word level such as *cheek bone*. Compounding is most often contrasted with overt affixation which derives a word from a lexeme by adding an affix, that is, a bound morpheme that combines with a specific category of base to form a pattern. An example of suffix derivation with a simple lexeme as a base is *wire+less*. A crucial feature of these combinatorial word-formation processes is that they are recursive and, as such, result in a hierarchical structure with binary groupings at each level of combination as the structures in (1) show:

- (1) a. [[[stress_N] [ful_A]] ness_N]
b. [[[smart_A] phone_N]] company_N]

Furthermore, compounds and affixations are morphosyntactically speaking headed structures. The suffix *-ful* in (1a) creates an adjective from the noun *stress*; the complex adjective *stressful* can serve as a base for further affixation by the suffix *-ness* which renders *stressfulness* a noun. Hence, each of these suffixes determines the word category of its derivative. Compounds are also headed in this structural sense. In the English example (1b), the head is the right-most constituent at each level of combination, but the head position can vary from language to language. The most productive compound patterns containing two native noun stems in the Romance languages, for example, have their heads on the left (cf. Rainer and Varela 1992, Scalise 1992, Fradin 2009, Kornfeld 2009). Consequently, affixation and compounding share most of their formal properties: they are binary branching, recursive, headed structures. Especially in languages that have right-headed compounds, like the Germanic languages, the primary difference between affixation and compounding lies in the status of the constituent parts: if at the relevant level of analysis both constituents are lexemes belonging to the open word classes of the language, the result is a compound, if one constituent is a formative, that is, a bound morpheme belonging to a finite class of elements in the language, the structure is an affixation.

Semantically the two types of construction tend to differ. An affix adds a general meaning component to its base. The suffix *-er*, for instance, denotes the agent of some activity, *-less* signals absence of some entity, *-ish* similarity with some property, the prefix *un-* negation of some feature, etc., so the affixations *dancer* ‘one who dances,’ *worthless* ‘without worth,’ *reddish* ‘slightly red’ and *untidy* ‘not tidy’ carry clear and explicit meanings. In a major class of compounds, often termed root or primary compounds, on the other hand, the connection between the denotation of the constituents is not overtly expressed: *Monsoon wedding*, *cadaver dog*, *sandwich war*, and *lawyer joke* are open in meaning until the intended relation is discovered. (Section 3.5 discusses a second large class of compounds, the verbal or synthetic compounds, whose interpretation is more specific in that it is based on the argument structure of the head.)

The notions free vs. bound form as well as that of a general meaning component can be quite elusive, however. Hence, obstacles arise in the demarcation of derivation from compounding when the decision as to whether a particular morpheme constitutes an independent lexeme, or whether it carries a generalized meaning, becomes hazy. This central problem is taken up in Section 3.2. Then Section 3.3 continues this discussion by dealing with the problem of bound roots, unique morphemes, neoclassical combining forms, and verbal prefixes and particles. Section 3.4 examines the interesting phenomenon of bound roots and lexical affixes in the incorporating languages. The structural ambiguity of the class of synthetic (or verbal) compounds is the topic of Section 3.5, and, finally, important ambiguities that arise between the products of compounding and other types of derivational processes such as conversion, back-formation, analogy, and different types of truncation that operate on complex bases as well as reduplication that creates a complex base constitute the topic of Section 3.6. Following these discussions a summary is given.

3.2 LEXEME OR AFFIX?

3.2.1 Transition from Lexeme to Suffix

A major problem in distinguishing derivation from compounding stems from the fact that—as the result of natural events occurring in the historical development of a language—an affix may emerge from an independent lexeme. To be more precise, Dalton-Puffer and Plag (2000) show that the development of the nominal suffix *-ful* in the Modern English pattern *cupful*, *handful*, *spoonful*, *mouthful*, etc., began in the 19th century on the basis of a phrasal structure in which a noun denoting a container functioned as the head of a complex noun phrase modified by an adjective phrase containing as its head the relational element *full*. Over the course of time, collocations such as *2 cups full of rice*, *3 barrels full of wine*, and the like underwent a series of interrelated developments: the plural marker on the nominal container began to shift to the end of the collocation, the spelling of the adjective *full* was reduced to *ful*, and the whole phrase came to be written as a complex word (i.e. *cupfuls*, *barrelfuls*). As a final result, the original adjective *full* had given way in this particular environment to a bound element *-ful* with nominal features. These changes from an independent adjective to a noun-creating formative are so radical that they leave little doubt that a new suffix pattern had emerged.

It is characteristic of the transition from an independent lexeme to a suffix for the lexeme to pass through a stage in which it is entrenched in a collocation and fixed in a specific order. A case in point is the Romance suffix—*ment(e)* that derives adverbs from adjectives (the ensuing discussion is based on Detges forthcoming) as in French: *lentement* ‘slowly’ < *lent*, *-e* ‘slow’; Italian: *chiaramente* ‘clearly’ < *chiaro*, *-a* ‘clear’; and Spanish: *generosamente* ‘generously’ < *generoso*, *-a* ‘generous.’ Historically, *-ment(e)* goes back to the ablative form of the feminine Latin noun *mens*, *mentis* ‘mental disposition, mind.’ As an independent noun in Classical Latin, it could be modified by an adjective phrase as in *mente valde placida* ‘with a very calm mind’ and alternate with other semantically similar head nouns in the same phrasal position such as *pectore* ‘breast,’ *corde* ‘heart,’ and *animo* ‘mind,’ for example, *laetanti pectore* ‘joyfully,’ *ardenti corde* ‘ardently,’ *studioso animo* ‘eagerly.’ With increased frequency, the *mente* construction became fixed in the order adjective + *mente* without intervening elements and, according to Detges, could at this stage (i.e. in the Classical Latin period before 200 AD) be considered a compound comprised of an adjective together with the noun *mente* because *mente* had not yet lost its nominal features. The transition from the head of a nominal compound to an adverbial suffix can be shown to have taken place when the construction shifted from its attitudinal meaning to a non-attitudinal one that could no longer be related to the “intention, disposition” meaning of the earlier nominal head of the compound. This stage is documented in the Reichenau Glosses from the 8th century where, for example, the word *solamente* is discussed as being in use in the spoken language in the same

meaning and function as the Classical Latin adverb *singulariter* ‘individually, one by one.’

The development from a compound constituent to a suffix is documented in the Germanic languages as well, cf., for example, for the German forms *-heit*, *-lich*, *-schaft*, *-sam*, *-tum* as well as for their English cognates. Henzen (1965: 110) observes that words whose meaning predisposes them to serve as elements of compounds may lose their independence in proportion to the productivity of the compound pattern of which they are a part. Erben (1983: 125–6) considers the grammaticalization from an independent word to a suffix to be complete when the original form no longer occurs independently, or at least when it can no longer be associated with the new form phonetically or semantically. For example, the Modern German suffix *-heit* stems from the Old High German noun *heid/heit* meaning ‘kind, appearance, status.’ In the 8th century compounds ending in *-heit*, such as *mana-heit*, *narra-heit*, are recorded and, around the year 870, twelve compounds ending in *-heit* are documented in Otfrid’s *Evangelienbuch*. Most of these compounds are formed with adjectival first constituents, for example *bōs-heit*, *kuon-heit*, *tumb-heit*, and serve as precursors for the New High German suffix pattern denoting abstract deadjectival nouns as in *Bos-heit* ‘meanness, malice,’ *Kühn-heit* ‘boldness,’ and *Dumm-heit* ‘dumbness, stupidity.’ Erben (1983: 127) gives an Old High German example in which the free form *heit* and a combined form *zága+heit* occur together in a single sentence. The gloss indicates the degree of meaning separation that distinguishes the two uses at this stage of the language [my emphasis, S.O.]:

- (2) *uuas nihein héit thúruh sina zágaheit*
 was no *personage* through his *timidness*
 ‘[he] was not a great personality due to his timidness’

By Middle High German times the independent noun *heit* was disappearing from the language as the growing number of combinations in *-heit* began to outnumber and overtake the older suffix pattern of abstract nouns ending in *-ī* (surviving into the modern language in forms such as *Dicht-e* ‘thickness,’ *Fläch-e* ‘flatness,’ *Näh-e* ‘closeness’). In Modern German, the suffix *-heit* has become the most productive formative in the creation of deadjectival abstract nouns and the noun *heit* no longer exists in the standard language. Erben attributes the success of the *-heit* pattern in suppressing the *-ī* pattern to the clearer structure of the *-heit* words at a time when the *-ī* suffix was undergoing a phonetic weakening that applied to all vowels in the final syllable of a word. In a like manner, the suffix *-lich* has its roots in compounds with Old High German *lih* ‘body’ as a second constituent, the suffix *-schaft* developed out of compounds with Old High German *scaf* ‘state, condition,’ *-sam* from compounds with Old High German *-sam* ‘same,’ and *-tum* from compounds with Old High German *tuom* ‘judgment’ (see Erben 1983: 126–8).

A similar genesis can be traced within the history of the English language in the case of the suffixes *-hood* and *-dom*. Modern English *-hood* arose from the Old English noun *hād* ‘state, rank, condition’ so that formations like *childhood*, *statehood*, *fatherhood*, etc., were originally compounds. And Modern English *-dom* developed out of Old English

dōm ‘judgment, law, state,’ cf. *freedom*, *wisdom*, which also took on the additional meaning of ‘territory’ in Middle English in words such as *kingdom*. Trips (2009) provides a detailed discussion of the history of these suffixes and Marchand (1969: ch. 4) sketches the earlier development of *-ly*, *-ship*, and *-some* into suffixes as well.

3.2.2 The Term Semi-suffix

Synchronically it is possible to observe patterns of formations that appear to be caught up in the transition from compounds to suffixations sketched in the previous section. For example, Marchand characterizes the elements *-monger*, *-wright*, and *-wise* (as in *warmonger*, *playwright*, and *crosswise*) as being “[h]alfway between second-words and suffixes.” These forms are no longer in use as independent words in Modern English; nevertheless, Marchand (1969: 210) argues that they are still “felt to be words” and therefore considers them semi-suffixes. Other examples seen by Marchand as belonging to the category semi-suffix are *-like* and *-worthy*. Although *manlike* appears upon first glance to be a compound made up of a noun and adjective, negated forms such as *ungentlemanlike*, *unbusinesslike*, *unsportsmanlike* show that *-like* formations have become reanalyzed as denominal suffixations that allow prefixation by means of the negative prefix *un-* which attaches to adjectives and adjectival derivations (but not to compounds). The same logic applies to *-worthy* formations, cf. *unpraiseworthy*, *untrustworthy*.

Fleischer and Barz (1995: 27) discuss the advantages of postulating an intermediate category for similar cases in German where a word appears both independently and in a series of formations. The primary motivation for a category semi-suffix (German *Halbsuffix*, *Suffixoid*) according to these authors is to be found in the weakening or generalization of meaning displayed by the proposed semi-suffixes vis-à-vis their independent counterparts, as well as in their characteristic distribution in a series of formations. Such criteria indicate that the combined form has distanced itself from its free variant and is possibly on its way to developing into a suffix. The authors are, however, in actual fact hesitant to accept such an intermediate category even though they acknowledge that phenomenon itself exists and in the 4th revised edition of their handbook—Fleischer and Barz (2012)—reject it altogether.

The German noun *Gut* ‘goods’ provides an example. Due to its relatively general meaning, it occurs in many combinations as a second constituent. In a number of these it yields a collective meaning “material needed for V” where a verbal first constituent provides information about the specific process involved: *Back-*, *Mahl-*, *Pflanz-*, *Streu-*, *Walz**gut* ‘material for baking, grinding, planting, spreading, crushing.’ With nominal first constituents that denote an abstract cognitive concept, a collective reading results that can be rendered as “N assets”: *Bildungs-*, *Gedanken-*, *Lied-*, *Kulturgut* ‘educational, thought, song, culture assets.’ As a result of the minor semantic distance between the *-gut* of the combined forms and the independent word *Gut*, Fleischer and Barz (1995: 143) consider these combinations compounds. Similarly, the relatively general German noun *Zeug* ‘stuff’ recurs as the second constituent in

combinations denoting “a group of utensils connected with a verbal activity”: *Ess-*, *Näh-*, *Rasier-*, *Schlag-*, *Strick-*, ‘eating, sewing, shaving, drumming, knitting utensils.’ Again, Fleischer and Barz (1995: 144) consider these constructions to be compounds. So here we find concord between Marchand (1969: 210) and Fleischer and Barz (1995) when the former argues that the fact that a word occurs frequently as a second element in combinations does not mean that it must have suffix status. As examples, Marchand cites English *proof* as in *bombproof*, *fireproof*, *rainproof*, *soundproof*, *waterproof*, and *-craft* as in *mothercraft*, *priestcraft*, and *witchcraft*.

Nevertheless, Fleischer and Barz (1995: 177–8) go on to classify combined forms ending in *-werk* and *-wesen* as suffixes. *Werk* as an independent noun means ‘work, production, opus.’ In combinations it may denote a work of nature as in *Ast-*, *Laub-*, *Buschwerk* ‘branches, foliage, shrubbery,’ artifacts made with a certain material, cf. *Leder-*, *Pelz-*, *Zuckerwerk* ‘leather, fur, sugar work,’ or collectives such as *Dach-*, *Balken-*, *Gitter-*, *Mauerwerk* ‘roofing, timberwork, grating, masonry.’ The noun *Wesen* has the meaning ‘essence, character, being.’ As the second element in a combination it takes on a more general meaning denoting the total collection of all offices and processes belonging to an institution: *Kredit-*, *Rechts-*, *Schul-*, *Gesundheits-*, *Finanz-*, *Strassen-*, *Versicherungswesen* ‘system of credit, law, school, healthcare, finance, traffic, insurance.’ Apparently Fleischer and Barz find the difference between “system of N” in the combinations and “essence, character” in the independent noun significant enough to merit the classification of *-wesen* as a suffix and similarly for *-werk* vs. *Werk*, although it is not clear why. *Laubwerk* and *Lederwerk* do not seem to be any less compound-like than *Nähzeug* and *Strickzeug*. Erben (1983: 81), on the other hand, considers all these formations, that is combinations in *-gut*, *-zeug*, *-werk*, and *-wesen*, semi-suffixes.

The conclusion, then, must be that the postulation of an intermediate category between a lexeme and an affix does not guarantee any real clarity in dealing with the question of the delineation of an affix from a lexeme, and thus serves no function. But upon closer examination, other problems accrue with the use of the term. Certain lexemes lend themselves easily to combinations in which they are specified via a co-constituent. The word *free* is a relational adjective and as such is easily combinable with its thematic object, both in phrasal constructions (*free of pain*, etc.) as well as at the word level, cf. *crisis-free*, *error-free*, *fat-free*, *pain-free*, *sugar-free*, *stress-free*, *tax-free*, *traffic-free*. These examples demonstrate that compounds group naturally around certain core lexemes into constituent families. The meaning of the core constituent in a constituent family may deviate from the central meaning of the independent lexeme. For instance, the compound *US-friendly* is understood literally as ‘friendly to/with the US,’ whereas *-friendly* in the combinations in (3) has shifted in meaning to signal ‘helpful, accommodating,’ a semantic extension associated with the central meaning of *friendly*, although not identical to it:

- (3) *user-friendly*, *reader-friendly*, *listener-friendly*, *environment-friendly*, *planet-friendly*, *industry-friendly*, *consumer-friendly*, *child-friendly*

Classifying *-friendly* as a semi-suffix on the basis of this meaning extension would characterize it as suffix-like in its properties and, in so doing, obscure an essential aspect of the nature of compounding.

3.2.3 “Morphological Transcendence”

Shifted meaning in combination with another lexeme is not specific to semi-suffixes, but is a more general phenomenon and is especially true of compounds. A novel compound must have a compositional meaning in order to be understood, but once a compound is accepted by a speech community it may take on idiosyncratic properties that result in the loss of its original transparency. The current consensus in psycholinguistics is that access to complex words in the mental lexicon proceeds via two different modes simultaneously—the parser automatically attempts to decompose the complex into its constituents while at the same time implementing a search for a whole-word entry, cf. the dual route models of Caramazza et al. (1985) and Frauenfelder and Schreuder (1992). In a series of psycholinguistic experiments, Libben (1994) provides additional evidence that the parser does indeed access all possible morphological analyses, a view also shared by, *inter alia*, Kuperman et al. (2010) and Ji et al. (2011). Using ambiguous novel compounds as stimuli, Libben forced his participants to decompose them by asking them to pronounce the words. *Busheater* and *seathorn* were read as *bush+eater* and *sea+thorn* rather than as *bus+heater* and *seat+horn*, a choice obviously influenced by the English digraphs <sh> and <th>. In a follow-up experiment, however, the reaction times required for a lexical decision on orthographically constrained ambiguous novel compounds such as these were the same as for orthographically unconstrained ambiguous novel compounds, for example *feedraft* (*fee+draft*, *feed+raft*). Both types of ambiguous novel compounds—orthographically constrained and unconstrained—required higher reaction times than unambiguous novel compounds such as *larkeater*. These results indicate that the orthographic constraints operate post-lexically, that is after all possible parses are generated. Furthermore, in the first test, no significant difference between the two possible parses for orthographically unconstrained ambiguous novel compounds was found, that is between *fee+draft* and *feed+raft*. The results did show, however, that there were stable preferences for one of the choices in each individual case that seemed to be based on semantic plausibility (for *cartrifle*, *cart+rifle* was spoken more often than *car+trifle*, but *car+driver* was chosen over *card+river* for *cardriver*). In order for a decision to involve semantic considerations, all parses must first be made available: as with the orthographic factor, a choice based on semantic plausibility must operate post-lexically. Libben et al. (1999) confirmed this finding in two further experiments by showing that ambiguous novel compounds prime associates of both possible parses. The stimulus *clamprod*, for example, primes both *sea* for *clam* and *hold* for *clamp*.

The activation of all possible parses is termed by Libben “maximization of opportunity.” The disadvantage incurred by the activation of all possible morphological analyses is that some of the activated information will be redundant. This disadvantage

is counterbalanced, however, by the need for the quick and efficient retrieval of meaning. The availability of all possible morphological analyses guarantees that no time-consuming reanalysis is required in case of an incorrect parse. Nevertheless, accessing a non-transparent compound under such conditions will result in a conflict between the whole word meaning and the meaning of the constituents. Tests show that exposure to a constituent prior to the presentation of a transparent compound facilitates access to the compound. Opaque compounds, on the other hand, cannot be primed in this way by their constituents, cf. Libben et al. (2003). This is known as the semantic transparency effect. Hence, the activation of, for example, *butter* and *fly* in addition to *butterfly* generates a conflict during parsing in need of resolution. At first it was believed that irrelevant information such as the meanings of non-transparent constituents could simply be suppressed. Due to findings in Libben (2010), the inhibition hypothesis has given way to the view that the mental lexicon is actually organized in a different manner. Rather than the deactivation of superfluous information, such conflicts cause the non-transparent constituents of opaque compounds to undergo a process of separation from their corresponding free form. This separation of meaning, termed by Libben “morphological transparency,” involves a semantic weakening or an increasing degree of abstraction such that the bound constituent transcends the meaning of its independent form. Hence, the process of lexical access induces compound constituents to establish their own positionally bound entries in the mental lexicon independent of the original free form whenever a conflict is perceived. The more often such a constituent is used as part of a compound, the stronger its representation will become and the less activation (and hence competition) will result from the free form. Evidence that this is the correct explanation for the semantic transparency effect is provided by lexical decision tests with words and non-word stimuli carried out by Nault and Libben (2004). Some of the non-words were lexemes that serve as the initial constituents of compounds. These resulted in greater rejection times as well as in a greater number of false positives than the non-constituent non-word stimuli did.

3.2.4 Essence of an Affix

The findings of the previous discussion lead us to assume that the mental lexicon contains, for example, in addition to an entry for the adjective *friendly*, an entry for the positionally bound constituent *-friendly* ‘helpful, accommodating’ found in the compounds in (3) which serves the purpose of alleviating direct competition with its free counterpart during access. A more complete understanding of the content and processes of the mental lexicon, therefore, sheds light on the natural process of meaning separation found in the case of compound constituents and calls into question the relevance of an intermediate category semi-suffix.

It is natural for speakers to construe compound constituents as bound variants of the corresponding free forms and to set up entries for them in their lexicon. The intuition that speakers have that allows them to differentiate between the constituent of a

compound and an affix arises on the basis of their implicit knowledge of the content of their mental lexicon. Marchand's hesitation to assume affix status for *-craft* in *witchcraft*, *priestcraft*, etc. (see Section 3.2.2) obviously has to do with the presence of the entry *craft* (as well as the related *handcraft*, *craftsman*, *crafty*, etc.) in the vocabulary. And *-monger* and *-proof* must still possess some degree of autonomy as a positionally bound noun and adjective, respectively, in the modern vocabulary to enable them to enter into new compounds such as *anger-mongers* or *rumor mongering* (*TIME* Feb. 18/Jan. 14, 2013) or appear as deadjectival converted verbs *climate-proofed* or *sound-proofed* (*The New Yorker* Jan. 7/March 4, 2013), cf. Section 3.5.1 where it is shown that suffixations as a rule do not undergo conversion. Thus, a deeper understanding of the nature of the compounding process speaks for a more perspicuous use of morphological categories. If the separation of meaning between a compound constituent and its corresponding free form is a natural phenomenon, the establishment of positionally bound compound constituents, and with them their constituent families, is not an indication of the beginning of a grammaticalization process leading to the emergence of an affix. This happens only under specific conditions. Consequently, the term affix should be reserved for reference to a pure formative, that is, a bound morpheme for which there is no competition with free lexeme in the mental lexicon, and the term semi-suffix is best avoided.

3.3 BOUND LEXEMES

3.3.1 Bound Roots, Unique Morphemes and Neoclassical Combining Forms

In spite of the courageous definition of the term affix just provided, one might wonder whether more needs to be said in order to distinguish an affix from a bound root. Bound roots are basic morphemes that have all the properties of lexemes except that they do not occur freely as, for example, *spec-* in *special*, *specific*, *specify*, *speciality*, and *ident-* in *identity*, *identical*, *identify* (Schmid 2011: 40). These words have been borrowed into English in their complex forms from Latin and French where they originated as derivations. But it is neither necessary to appeal to this knowledge (which many speakers lack anyway) nor to the higher degree of lexical-semantic content characteristic of bound roots vs. the more abstract semantics of affixes to differentiate the two. Apart from their distinct phonological differences from affixes, bound roots cannot be affixes because they co-occur with affixes and by definition affix + affix combinations are not possible.

Unique bound forms are roots that only occur once in the vocabulary such as the underlined portion of English *unkempt* or of German *Unflat* 'filth.' Unique bound forms are not restricted to occurring only as bases in combinations with affixes, as are the bound roots discussed in the previous paragraph, but are also found in combinations with stems, cf. English *raspberry*, *lukewarm*, and *nightmare* and German *Schornstein* 'chimney' and *Bräutigam* 'bride groom.' Hence, the argument against affix +

affix combinations just given cannot exclude them from being affixes. However, unique forms are one-time occurrences and thus differ markedly from affixes, which recur in a series of combinations. Furthermore, knowledge of the content of the mental lexicon will include the awareness of a closed class of affixal formatives with their characteristic phonological properties. The second element in *doughnut*, for example, would not be perceived as a suffix, whereas the ending of *laughter* might. Although not productive, (the remnants of) a pattern could be surmised for *-ter* on the basis of its similarity with *slaughter* since both words consist of a verbal stem and have an event/result meaning related to that stem. In addition, the monosyllabic form of *-ter* which, in contrast to *-nut*, contains a reduced vowel is not a possible stem.

A different sort of bound root is found in the neoclassical compounds that are prevalent in most of the modern languages of Europe. Neoclassical compounds are combinations of Greek and Latin lexemes that are formed according to the compounding rules of the modern languages, cf. English *neurology*, *democracy*, *stethoscope*, *suicide*, *anglophile* (Bauer 1998), French *aérodrome*, *hiéroglyphe*, *géographe*, *anthropomorphe*, *hétérodoxe*, *pathogène* (Zwanenburg 1992), Polish *fotografia*, *makroekonomia*, *neofita* 'neophyte', *poligamia*, *kse-nofobia* 'xenophobia', *neurologia* (Szymanek 2009) and Basque *telefono*, *mikrobiologia*, *filologia*, *elektromagnetismo* (Artiagoitia et al. forthcoming). The combining forms used to create neoclassical compounds do not occur as independent words in the modern languages and are, furthermore, are often restricted to either the initial or final position of a combination; for example, *astro-*, *bio-*, *electro-*, *geo-*, *gastro-*, *tele-* occur initially while *-cide*, *-cracy*, *-graphy*, *-phobe*, *-scope* occur finally (cf. Plag 2003: 155ff.). Hence, as Bauer (2005a) states, neoclassical compounds do not fit the definition of compounds and this is precisely the motivation for establishing a special category to accommodate them. In their formal aspects they even seem to have much in common with prefixes and suffixes. However, prefixes and suffixes do not combine with one another as the neoclassical combining forms characteristically do. So the need does not arise to appeal to the fact that they were lexemes in their source language to exclude them as affixes in the modern languages.

The establishment of a special category of bound forms is also the best course of action for a problem to which Aronoff (1976) drew attention, namely the case of the Latinate verbs in English whose structural components are not morphemic in the strict sense as in *permit*, *remit*, *submit*, *transmit* or *conceive*, *deceive*, *conceive*, *receive*. The bound units in these structures differ from the neoclassical forms in that they are without an identifiable component of meaning and, hence, do not function as combining forms. Nevertheless such words are analyzable as containing two structural units as the regular allomorphic variation of their stem demonstrates, cf. *permission*, *submission*, *transmission* and *conception*, *deception*, *reception*, etc.

The neoclassical combining forms, on the other hand, are productive elements that are not necessarily restricted to only combining with one another—many of them also combine with native roots in the respective language as, for example, English *speed+ometer*, *mob+ocracy*, *Kremlin+ology*, *weed+icide*, *chimp+onaut* (Adams 1973, 2001, Bauer 1998) and Polish *fotokomórka* 'photocell', *kryptopodatek* 'crypto-tax', *pseudokibic* 'pseudo-fan', and *helmofoon* 'headset' (Szymanek 2009). In these combinations, they can give rise to

new constituent families. Formations like the German *Kartothek*, *Filmothek*, *Spielothek* ‘collection of maps, films, games’ in addition to *Bibliothek* ‘archive of books’ demonstrate this. A particularly interesting example of relevance for the discernment of derivation from compounding concerns the final combining form *-itis* which signals ‘disease, inflammation’ in combination with an initial combining form, cf. *appendicitis* ‘inflammation of the appendix,’ etc., in (4a). However, when *-itis* appears with a native English lexeme, its meaning shifts to ‘addiction, abnormal excess of,’ cf. (4b).

- (4) a. appendicitis, arthritis, encephalitis, gastritis, laryngitis, meningitis,
tonsillitis
b. computer+itis, cellphone+itis, facebook+itis, junk-food+itis,
telephone+itis

So *-itis*₂ of (4b) has established itself as a second bound form in an extended, but related, sense to the combining form *-itis*₁ in (4a). In addition, *-itis*₂ displays a phonological form that is quite similar to other suffixes in English; it begins with a vowel (as do, e.g., *-ion*, *-ic*, *-ify*, *-ize*) and is bisyllabic with a strong–weak stress pattern that conditions a base ending in a weak stress and, thus, consists of at least two syllables. These are properties that are typical of suffixes, in particular the type of suffixes that have been termed “non-stress neutral” or were thought to belong to class 1 in Level Ordering theories such as Allen (1978), Siegel (1979), and Kiparsky (1982b), and hence make *-itis*₂ (lacking a free counterpart in the English lexicon) quite suffix-like. Factors ruling against this characterization are the existence of the related neoclassical pattern and the relatively restricted number of formations compared to more typical cases of affixation.

3. 3.2 Prefix vs. Preposition and Adverb

Traditional grammars have a history of treating prefixation, not together with suffixation as a type of derivation as modern linguistic theory does, but as a type of compounding. For the Germanic languages, this was the case *inter alia* in Herman Paul’s (1955) *Deutsche Grammatik* as well as in the first edition of Walter Henzen’s (1947) *Deutsche Wortbildung*. Bauer (2005a) reports that this tradition was prevalent in Romance linguistics as well. The reason for this was the historical awareness that many prefixes originated in prepositions and adverbs that occurred as first constituents of compounds, accounting for their function as modifiers of the head element rather than as formatives for new words like suffixes. The transition from free prepositions/adverbs to first forms of compounds and finally to bound formatives follows a path similar to that sketched in Section 3.2 for suffixes. Prefixation is exceptionally productive in the formation of verbs. The verbal prefixes in Modern German *be-*, *ent-*, *er-*, and *ver-* derive from earlier prepositions but no longer have free counterparts. They are unstressed and inseparable from their stems, cf. *sie besprechen das Band* ‘they are recording the tape.’ Prepositions, in their intransitive use as adverbs, often appear together with a verb stem as particle

(also termed phrasal, multi-word or compound verbs). The prepositions heading the PP complements in (5a and b), for example, take an NP object. In (5a' and b') the same forms occur intransitively as particles forming the complex verbs *aufsprühen* and *aus-schütten* where the particles *auf* and *aus* are stressed and occur separately from the verb stem in all clauses that require the finite verb to occur in the second (functional) position of the clause. The same phenomenon, including the separation of the particle from its verb stem, occurs in English as can be seen in the glosses in (5), and indeed this phenomenon is found throughout all the Germanic languages.

(5)	a.	Sie sprüht die Farbe <u>auf</u> die Wand	'She sprayed the paint <u>on</u> the wall'
	a'.	Sie sprüht die Farbe <u>auf</u>	'She sprayed the paint <u>on</u> '
	b.	Er schüttete das Wasser <u>aus</u> dem Glas	'He poured the water <u>out</u> of the glass'
	b'.	Er schüttete das Wasser <u>aus</u>	'He poured water <u>out</u> '

Interestingly, there is a small class of prepositions in Modern German that allow both intransitive particles and also have prefix variants. The contrast between the two constructions clearly demonstrates the difference between a verbal particle and a verbal prefix. The prepositions *durch* 'through', *über* 'over', *um* 'around', and *unter* 'under' belong here; they occur as separable verbal particles as the examples in (6a' and b') show and as prefixes as in (6a'' and b''):

(6)	a.	Die Mücken fliegen um die Kérze	'The gnats are flying around the candle'
	a.	Die Mücken fliegen <u>um</u>	'The gnats are flying around'
	a''.	Die Mücken <u>um</u> fliegen die Kerze	'The gnats are flying around the candle'
	b.	Die Bande streift durch die Stádt	'The gang is roaming through the city'
	b'.	Die Bande streift <u>durch</u>	'The gang is roaming through'
	b''.	Die Bande <u>durch</u> streift die Stadt	'The gang is roaming through the city'

The difference between a particle and a prefix is found in the separability and stress on the particle vs. the inseparability and lack of stress on the prefix. Moreover, the particle defocuses the prepositional object by suppressing it formally; hence, its existence is implicit and presupposed. The prefix verb, on the other hand, inherits the original object of the preposition that is incorporated into the verb stem and expresses it as its own direct object, cf. Olsen (1996).

3.4 BOUND ROOTS AND LEXICAL AFFIXES IN THE INCORPORATING LANGUAGES

Bound roots discussed in Section 3.3.1, which appear to be a relatively marginal phenomenon in the European languages, constitute a more regular phenomenon in other

languages such as the noun incorporating languages. An example of noun incorporation is the complex verb in (7a), taken from the Iroquoian language Tuscarora, which contains the verb root *-ù:rəl-* ‘split’ and an incorporated the noun root *-rəlʔn-* ‘log.’ Roots in the Iroquoian languages are bound; they must combine with affixes to occur freely. The verb root in (7a) occurs with three prefixes, the first two are inflectional and the third is pronominal, expressing first person singular and satisfying the external argument of the verb, cf. Mithun (2000: 916).

- (7) a. /waʔ-t-k-rəʔn-ù:rəʔ-ʔ/
 AOR-DU-1.SG-log-split-PFV
 ‘I split a / the logs’
- b. /u-rəʔʔn-e waʔ-t-k-ù:rəʔ-ʔ/
 N-log-NOM.SUFF AOR-DU-1.SG-split-PFV
 ‘I split a / the logs’

The incorporated construction typically has a counterpart in which the noun is found external to the verb. This is shown in (7b) where the verb root appears with its prefixes but without the nominal root. Instead, the nominal root, now marked with a neuter prefix and a nominal suffix, heads an independent noun phrase. While the independent noun constitutes its own phrase in the analytic construction occurring with functional elements that determine its reference, definiteness, quantity, etc., the incorporated noun root is devoid of such syntactically relevant markers and is understood generically, that is, as a modifier that restricts the type of activity denoted by the verb. The verbal meaning is narrowed from “splitting” in (7b) to “log-splitting” in (7a). This formal difference between the analytic and synthetic constructions spawns different functional uses. Pragmatically, the independent noun is used to introduce new discourse entities, to express focus or contrast and to signal salience; the incorporated noun root is chosen when the entity in its denotation is already present in the discourse or otherwise backgrounded. Incorporation often affects the verb grammatically as well, resulting in verbal diatheses such as intransitivization, passivization, and causativization (Mithun 2000). If compounding is defined as the combination of two lexemes (roots or stems), the complex verbs resulting from noun incorporation obviously qualify as compounds with bound lexical constituents.

The interest of such formations to the topic at hand, that is, the delineation of derivation from compounding, lies in their close relationship to a construction similar to noun incorporation but differing from it in that one of the constituents is formally an affix rather than a bound root. Nevertheless, the root+affix combinations in question share the grammatical, discourse, and semantic properties that are typical of regular noun incorporation: they result in verbal diathesis, carry distinct discourse functions, and derive subordinate level concepts prone to lexicalization. At the same time, the verbalizing (or nominalizing) affixes in question have meanings more typical of roots than of derivational affixes. Hence, these formatives have been termed “lexical affixes,” cf. Mithun (1999, 2000).

Mithun (1999) reports that numerous examples of lexical suffixes can be found in Spokane, a Salish language spoken in Washington State. The suffix *-cin* ‘mouth, food’ occurs in the complex noun *n-č’ m-cín* ‘the mouth of a river, lit. LOCATIVE-river-mouth.’ Yup’ik, an Eskimoan language spoken in Central Alaska, has an even larger number of lexical suffixes as, for example, the verbal *-cur-* ‘hunt’ in *nayircurtuq* ‘he is seal-hunting; lit. seal-hunt-INDIC.INTR-3SG.’ There are strong arguments, that in spite of their root-like lexical meanings, these units are indeed suffixes (cf. Mithun 1999: 48–56 and 2000: 922–3): In Spokane nominal roots can occur alone and in Yup’ik verb roots can appear with just an inflectional suffix, but in neither language is this possible for the lexical suffixes. In Yup’ik words, an initial root is followed by a series of suffixes but the lexical suffixes cannot assume the initial position; they must follow a root. Generally, the lexical suffixes are not cognate with any of the roots in the languages that have them, but they do have unrelated counterparts that occur as an independent root or stem under the discourse conditions that require the analytic rather than the incorporated version of the sentence. Although the semantics of the lexical suffixes is root-like, they tend to have a broad range of meanings that are typically more diverse than the meaning of equivalent roots. Derivations resulting from lexical affixes are often lexicalized and speakers are aware of which combinations exist and are in use in the speech community. Finally, the lexical suffixes serve as formatives in the creation of new lexemes: they recur in patterns much like derivational suffixes as can be seen in the Yup’ik pattern formed on the basis of the suffix *-cur-* ‘hunt’: *nayircurtuq*, *kanaqlaggsurtuq*, *tuntussurtu*, *yaqulegcurtuq*, *neqsurtuq*, *kayangussurtuq*, etc. ‘(he is) seal-hunting, muskrat-hunting, caribou-hunting, bird-hunting, fish-hunting, egg-hunting, etc.’ and speakers create new formations on the basis of such patterns (Mithun 1999: 51, 55). Furthermore, the individual lexical suffixes differ in their productivity, some being more productive than others, which is also a feature typical of derivational patterns. Moreover, lexical prefixes have been found in two Salishan languages, Bella Coola and Nuxalk, as well as in Nisgha, a Tsimshianic language, also of British Columbia (Mithun 1998: 300–1). Finally, even though the inventory of the lexical affixes in the languages that have them is quite large, they represent closed classes.

What is the explanation for this interesting mix of root-like semantics and affixal form? Mithun (1998, 1999, 2000) argues that the root-like meanings of the lexical affixes stem from their diachronic origin as roots of incorporated structures. The lexical affixes are found in semantic constellations typical of incorporated structures including the classificatory pattern exemplified in (8) by Mohawk, an incorporating language of the Iroquoian family, where the verb has incorporated a general nominal root “liquid” that is further specified by the more specific independent noun “milk,” rendering a meaning like “I liquid-consumed milk.”

- (8) Mohawk
 onòn-ta’ ‘wa’khneki-ra’
 milk I-liquid-consumed = ‘I drank milk’

Bella Coola has lexical suffixes that act similarly. In (9) the verb is comprised of the verbal root *-q̄is-* ‘scorch’ together with the nominal lexical suffix *-uc-* ‘mouth, food’ meaning ‘to cook; lit. scorch food.’ The independent object “strips of spring salmon skin” then specifies the type of food cooked (Mithun 1998: 306).

(9) Bella Coola

s-íq̄-kʷ	ta-s-q̄is-uc-im-tx
NOM-split-QUOT	PROX-NOM-scorch-food-PASSIVE-ARTICLE
‘What he cooked was strips of spring salmon skin’	

The Salishan languages are verb initial and many of them still have verb-noun incorporation structures in addition to lexical affixes. Interestingly, the lexical suffixes are nominal and the lexical prefixes are verbal (although the lexical affixes are not cognate with the roots), cf. Mithun (1998: 308). Although the present-day Eskimo-Aleut languages—as opposed to the Salishan languages—do not allow noun incorporation in its root+root form, Mithun (1998, 1999, 2000) assumes that it may have been an earlier option and, again, the source of the lexical affixes. If the lexical affixes have indeed developed via a grammaticalization process from the lexical roots of earlier noun incorporation structures, this would explain their relatively large number, since their source would have originated in the open classes of verb and noun roots. It also explains their root-like semantics as well as their slightly more generalized meaning. Mithun (2000: 926) sketches a plausible scenario for this transition from bound lexical root to lexical affix: Assuming that the productivity of noun incorporation in the source language begins to diminish, over time two types of language change could obscure the relation between the original constituent roots of the incorporation structures and their independent counterparts, inducing a gradual grammaticalization process that could result in lexical affixes: First, incorporation structures that had already become an established part of the vocabulary would continue on their own individual course of development as autonomous words, allowing more general and diverse semantic aspects to creep into the constituent parts while still maintaining the basic core of their original meaning. Second, individual roots in the vocabulary could fall out of use in time. In this case, their incorporated counterparts would be prone to reanalysis as formatives without independent counterparts. If this scenario is correct, the development of lexical affixes may well have proceeded in much the same way as sketched in Section 3.2 for the suffixes and prefixes of Romance and Germanic, that have also originated in independent lexemes of the open classes. Hence, it is worth asking whether the reason for the discrepancy in number of lexeme-to-affix cases in the non-incorporating vs. incorporating languages may lie in the nature of the lexeme that serves as their source. In the Romance and Germanic cases discussed above, the prerequisite for the grammaticalization of a lexeme as a derivational affix (suffix or prefix) was an intermediate stage as a compound constituent which—according to the discussion in Section 3.2.3—is encoded by speakers as an additional entry in the mental lexicon for a positionally bound variant of the free lexeme. Incorporating languages, in which the lexemes are bound roots from the

start, satisfy this precondition generally, allowing the reanalysis to draw from a larger pool of lexical sources.

3.5 SYNTHETIC COMPOUNDS

The term synthetic compound was apparently first coined by von Schroeder (1851–1920) for a class of complex words, exemplified by German *Machthaber* “power holder,” which appear to involve a synthesis of two formation processes: the first and second elements form a compound (cf. *macht+hab-*), while the second and third exemplify a derivation (i.e. *hab+er*). The peculiarity of this class of formations was that neither the first two elements alone nor the final two formed a word—a word only came about when all three elements occurred together. Wilmanns (1986: 2–3) referred to such constructions as *Zusammenbildungen* “together formations,” a term which has survived in its original sense in contemporary German linguistics and is employed in order to avoid the use of the term “compound,” cf. Neef (forthcoming). The issue from the start has been the uncertainty as to whether complex words like *power holder* should be analyzed as compounds (= (10a)) or derivations (= (10b)):

- (10) a. [[power]_N [hold_V+er]_N]_N
 b. [[power_N + hold_V]_V -er]_N

In English, the term synthetic compound is first found in Bloomfield (1933), where two types are distinguished, the denominal “synthetic” constructions such as *long-tailed* and the deverbal “semi-synthetic” constructions as in *meat eater* (ten Hacken 2010c: 233). Marchand (1969: 15–18) makes use of the term in a like fashion; the key words he uses as representative of the category are *watchmaker* and *hunchbacked*. Although Marchand explained the semantic properties of synthetic compounds by relating them to an underlying verbal nexus, he considered them to be genuine compounds that do not differ “at the level of morphologic structure” from regular primary compounds consisting of a N+N or a A+N as found in *steam boat* and *color blind*. Allen (1978), too, analyzed synthetic compounds as the adjunction of two lexemes with the same structure as primary compounds. It was most likely early insights like Marchand’s and Allen’s that opened the way for modern theories of word formation to extend the content of the term from its originally narrower sense to become equated with the class of compounds based on a verbal interpretation as a whole, that is, those whose interpretations arise on the basis of the argument structure of the deverbal head regardless of whether the head can occur alone (*programmer*—*computer programmer*) or not (??*keeper*—*house keeper*).

Some early approaches, however, were hesitant to assign synthetic compounds the same structure as primary compounds. Botha (1981), for example, argued extensively against this step for both Afrikaans and English. His analysis in which they were derived via affixation from underlying phrases ran into criticism because the A+N constituents

(cf. Dutch *blauwogig*, German *blauäugig*, and English *blue-eyed*) lack the inflection required in an NP, hence, therefore could not be syntactic phrases (cf. Booij 2002: 158, ten Hacken 2010c: 235). In accordance with her argument linking hypothesis, Lieber (1983) adopted the structure in (10b) for *truck driver* and assumed that *truck* was linked as an argument to the verb *drive* which headed the initial V constituent and formed the base of a derivation by means of the suffix *-er*. This solution incurred the objection that compound verbs of the form N+V (cf. **to truck-drive*) are not possible in English and hence do not constitute plausible bases for the very productive pattern of synthetic compounds, cf. Booij (1988: 67). In addition to this objection, Booij provided a counterargument against the structure (10b): the head of Dutch *aardappelgevrete* ‘excessive eating of potatoes’ is *gevrete*, a deverbal prefixation showing that *aardappel* does not constitute a constituent together with the verbal stem *vrete*, but rather the deverbal nominal head *gevrete* is first derived via prefixation at which point *aardappel* is adjoined via composition.

The majority of authors (Allen 1978, Selkirk 1982, Plag 2003, ten Hacken 2010c, among many others) have opted for the N+N adjunction shown in (10a) augmented by the concept of argument inheritance. The deverbal noun *holder* inherits a modified version of the argument structure of the transitive verb *hold* and assigns the inherited internal object to its first constituent (cf. Booij 1988, 2002, Lieber 2004, Jackendoff 2009). Although the internal arguments of nouns are in general optional, nouns derived from transitive verbs often sound odd when they occur alone, but their status improves when they appear with their internal object, cf. ??*installer* vs. *window installer*, *installer of windows*. When the object of the verbal activity can be inferred, the deverbal agent noun is often lexicalized, as in *settler* ‘homesteader’. The lexicalization process is independent of the productive word-formation pattern which draws on the lexical meaning of the constituents in a compositional manner just as the syntax does, cf. *score settler*, *Crowd pleaser*, *decision maker*, *page turner*, *blowout preventer*, etc., all fit this pattern. Hence, *pleaser*, *maker*, *turner*, *preventer* can be considered possible (relational) nouns that are most sensible when their meaning is completed by their objects. Although the most frequent affixes involved in synthetic compounding are *-er*, *-ing*, and *-en*, most linguists have followed Allen (1978), Botha (1981), Selkirk (1982), and others in assuming that the head constituent of a synthetic compound in this newer sense can arise by means of a wide variety of deverbal suffixes (cf. *globe-spanning*, *law enforcement*, *cost reduction*, *slum clearance*, *snow removal*, *teacher trainable*) as well as by conversion, cf. *tax cut*. In this context, Lieber (2010b) discusses a major difference in compounds with affixed and converted heads. The latter induce a verbal interpretation based on the subject argument much more freely than the former, cf. *fleabite*, *cloudburst*, *dogfight*, *footstep*, *heartbeat*, *sunrise*.

Interestingly, some linguists who consider the deverbal synthetic pattern *power holder*, *truck driver*, etc. to be compounds structured along the lines of (10a) still reject subsuming the denominal pattern *blue-eyed*, *open-minded*, *three-wheeler* under the same analysis. The stumbling block is the oddness of the derived head ??*eyed*, ??*minded*, etc. Plag (2003: 153) analyzes *blue-eyed* as a phrasal affixation in which the suffix *-ed* attaches to the noun phrase *blue eyes*, thus directly mirroring its semantics ‘having blue eyes’. Adhering to the same goal of proposing an analysis that directly accounts for the semantics of the

construction, ten Hacken (2010c: 240) adopts a structure along the lines of (10b) for *blue-eyed* and *three-wheeler*. But since the first constituent lacks the obligatory inflection that is required in a phrase (*blue eyes*, *three-wheels*), he assigns the A+N constituent the status of a “morphological phrase,” stipulating that morphological phrases are created in the lexicon which enables them to take part in word-formation rules, but do not enter the syntax.

Misgivings about the well-formedness of the simple derived heads that form a constituent under the analysis in (10a) need not stand in the way of a uniform treatment of the verbal-based and nominal-based synthetic patterns. Whereas the incompleteness of ??holder vis-à-vis *power holder* is related to the underlying transitivity of the verbal argument structure, the incompleteness of ??eyed, ??legged, ??wheeler, etc., could well be a matter of pragmatics. The simple adjective *bearded* is an acceptable word because not every man wears a beard and so the property of having a beard is informative. Adjectives expressing properties that all objects of a category automatically possess are not informative unless they contain a further specification that increases their information content as seen in *blue-eyed*, *short-legged*, *three-wheeler*, cf. Booij (2002: 158) and Neef (forthcoming). Consequently, a uniform analysis treating both variants of the synthetic pattern as regular compounds is both empirically sound and theoretically explanatory. Synthetic compounds with a deverbal head differ from primary compounds in that the relation upon which their interpretation depends is not implicit, but inherent to the verb and inherited from the verbal base by the derived head. In the case of synthetic compounds with denominal heads, the suffix expresses a possessive relation explicitly. Here there is no difference to normal compounds. If a derivative is uninformative (cf. ??haired) a compound structure has the potential of providing it with more information (i.e. *long-haired*). Compounds as well are subject to the same pragmatic requirement on informativeness, cf. ??page book, ??horn cattle whose status improves with more information: *200-page book*, *long-horn cattle*. Other views at odds with this conclusion exist. One example is Booij’s (2009: 212–14) analysis of synthetic compounds in construction grammar that unifies two schemata $[NV]_V$ and $[V\ er]_N$ into a single complex schema $[[NV]_V\ er]_N$. This solution formalizes the construction in more or less the descriptive terms of traditional grammar as recounted above, but offers no explanation for the extreme productivity of synthetic compounds in spite of the limited productivity of the proposed internal NV-constituent.

3.6 DERIVATION BY CONVERSION, BACK-FORMATION, ANALOGY, BLENDING, AND REDUPLICATION

3.6.1 Conversion

Conversion is the process by which a lexeme belonging to one lexical category is taken over into another lexical category; consequently, the lexemes in a conversion relation

share a phonological form and are closely related in meaning. The major patterns are verb to noun (e.g. *to fight* > *a fight*) and noun/adjective to verb (cf. *a text* > *to text* and *obscure* > *to obscure*). Conversion from suffixed bases is not usual. Marchand (1969: 372) sees the reason for this in the function of suffixes as categorizers: suffixes determine the category of a derived word and subsequent conversion, that is, an unmarked change in category, would obscure this function, cf. *happiness* > **to happiness*. No such obstacle stands in the way of converting a compound, however, because the head of a compound is not an affix but a lexeme. The major stock of compounds in a language is found in the open class of nouns, and compound nouns can indeed be converted to verbs, cf. *to blackmail*, *blindfold*, *broadside*, *earmark*, *handcuff*, *honeymoon*, *shortlist*, *skyrocket*, *spotlight*, etc. A new verbal form such as *to instant-message* must be seen as a product of the derivational process of conversion in the same way that *to text* < *a text* is and, hence, a derivative of the compound noun *instant-message*. It is important to see that the verb *to instant-message* is not a genuine compound. As such it would have to have been created by the free combination of the two lexemes involved, the noun *instant* and the verb *to message*. It is rather the case that the compounding process creates the complex noun *instant-message* which can subsequently be converted as a whole to the verb *to instant-message* with the related meaning 'to send an instant message.' Compound adjectives are not as frequent, but they can serve as the source of new verbs, cf. *soundproof* > *to soundproof* and *climate-proof* > *to climateproof* (as discussed in Section 3.2.4).

Conversion may interact with compounding on different levels. The result in each case is a slight addition in meaning that clearly reveals the derivational history of the complex word. For example, from the nominal compound *whitewash* denoting 'a liquid for whitening' the denominal verb *to whitewash* 'to whiten with whitewash, to gloss over or cover up' can be derived via conversion. This complex verb can then give rise to the converted deverbal noun *a whitewash* meaning "an instance of whitewashing" or "a cover-up." Hence, the noun *a whitewash* in this second sense is a derivation, while the original complex mass noun on which it is based, that is, *whitewash*, is a compound. The original nominal compound in fact consists of the two lexemes *white* and *wash*; the nominal *wash* meaning 'a liquid for washing' is a conversion from the verb *to wash*. The complex lexeme *a whitewash*, thus, has the derivational history shown in (11):

(11)	<i>to wash</i> > <i>a wash</i>	V conversion to N	'liquid for washing'
	<i>white</i> + <i>wash</i>	compound A+N	'liquid for whitening'
	<i>to whitewash</i>	N conversion to V	'to whiten with whitewash, to cover up'
	<i>a whitewash</i>	V conversion to N	'an instance of whitening with whitewash, an instance of covering up'

Consequently, one can give the side of a building a "wash" and one can give it a "white-wash" and one can also give something a "whitewash" in the extended sense of a cover-up, but *a whitewash* is not a result of the free combination of *white* and *wash*, it is a conversion from the verb *to whitewash*, just as *a wash* is a conversion from the verb *to*

wash. If the semantics of a *whitewash* relates it to the verb *to whitewash*, then it is a product of conversion and not of compounding.

3.6.2 Back-formation

Back-formation is a derivational process that is, in effect, the converse of affixation. Instead of adding an affix to a lexeme to derive a new word, back-formation creates a new lexeme via the subtraction of a supposed affix from an apparently complex base. The mistaken analysis is motivated by the phonological and semantic similarity of the supposed affixation to other cases of genuine affixation. Examples of new lexemes that arise via back-formation are *to cohes* < *cohesion* (cf. *act* > *act+ion*), *to seize* < *seizure* (cf. *close* > *closure*) and (a) *laze* < *lazy* (cf. *crisp* > *crisp+y*). Affixation relates a less complex word to a more complex one, while the direction of the relation is reversed in back-formation.

Synthetic compounds are a major source of novel back-formed verbs, cf. *chain-smoker* > *to chain-smoke*, *babysitter* > *to babysit*. Adams (2001: 118) points out that neoclassical compounds can also be back-formed to verbs, cf. *to biodegrade* < *biodegradable*. A novel back-formed verb—like *to anger-manage* from *anger management*—is not a genuine compound. A compound is a free combination of two lexemes with an open relation between them. A back-formed verb is a reduction of an existing complex word, hence its semantics will be based strictly on the meaning of the complex word; the constituents are not freely related to one another as in a compound. Therefore, *to anger-manage* cannot mean ‘to manage by the use of anger, in the form of anger, in spite of anger’ or any of the other possibilities which would be possible for a genuine compound. Its interpretation is in fact based on the motivating synthetic compound *anger management* and, hence, is restricted to the meaning ‘to manage anger.’ The same applies to the compound bases *chain-smoker*, *window-shopper*, and *ghostwriter* giving rise to the back-formed verbs *to chain-smoke*, *window-shop*, and *ghostwrite* with the non-compositional meanings denoting the activity of a *chain-smoker*, *window-shopper*, and *ghost-writer*.

3.6.3 Analogy

The process of analogy allows a new word to be created by analyzing a base as a formal and semantic complex A + B. If one of the elements, A or B, is exchanged for an element C, perceived as more appropriate for the desired meaning of the new word, either C + B or A + C arises. Examples of individual analogical derivations are *whitemail* < *blackmail*, *slow food* < *fast food*, and *underwhelmed* < *overwhelmed*. The complex words *whitemail*, *slow food*, and *underwhelmed* are not products of the process of compounding as seen by the restriction of their meaning to a variation of the meaning of the analogical base. *Whitemail* denotes the opposite of the process

of extortion encoded in *blackmail*. *Slow food* and *underwhelmed* are understood in opposition to *fast food* and *overwhelmed*. The openness of meaning characteristic of the compounding process is lacking.

The same reasoning applies when a series of formations develops on the basis of an original analogic formation. In (12a) the meanings of *First Couple* and the additional forms arise on the basis of the knowledge of the whole word *First Lady*. It is not the case that *First* alone takes on the meaning 'presidential.' Similarly, *landscape* in (12b) serves as the analogical basis for the other forms which include 'landscape' in their meaning, that is *moonscape* 'landscape on the moon.'

- (12) a. First Lady > First Couple, First Daughters, First Dog
 b. landscape > moonscape, seascape, cityscape, dreamscape, spacescape, streetscape

Each analogical series must be considered on its own merits; the words in (13), namely, are compounds containing the positionally bound constituent *e*, an abbreviated form of "electronic." This meaning component enters into the complex meaning of each of these words in a compositional manner, hence these formations are best considered as compounds.

- (13) e-mail > e-commerce, e-shopping, e-cash, e-business, e-delivery, e-readers

3.6.4 Blending

The process of blending has much in common with compounding, although it differs from compounding in its intentional nature. In blending, two lexemes are combined, but at the same time they are superimposed upon one another leading to a shortening of one or both constituents. Nevertheless, the meaning of each constituent lexeme flows into the meaning of the blend in the same manner as with compounds: *gundamentalist*, *screenager*, *stalkarazzi* are the equivalent of determinative compounds ('gun fundamentalist, i.e. fundamentalist with respect to guns,' etc.) while *dramedy*, *Spanglish*, *kidults* are the equivalent of coordinative compounds ('drama-comedy, i.e. both drama and comedy,' etc.). The shortened forms of the blend's constituents are subject to prosodic factors, which is not characteristic of compounding but typical of certain derivational processes. Plag (2003: 125) shows how the whole blend has the same number of syllables as the full form of the underlying second constituent:

- (14) a. *globe*(1) + *obesity*(4) = *globesity*(4)
 b. *guess*(1) + *estimate*(3) = *guesstimate*(3)
 c. *friends*(1) + *enemies*(3) = *frenemies*(3)

In spite of their reduced phonological form, the constituents of the blend retain the meaning of the original lexemes; in order to understand *threepeat* and *Twitizens* both components must be reconstructed: *three* + *repeat* and *Twitter* + *citizens*. A blend can serve as an analogical base, inducing further formations: *chocolate* + *alcoholic* has led to *chocoholic* which in turn has enabled, *carbaholic*, *workaholic*, *shopaholic*, *spendaholic*, and possibly others. Here the status of *-aholic* is more difficult. It hinges on whether it has come to mean ‘addicted’ on its own and can enter new combinations freely. But as long as its full form needs to be recovered to reconstruct the meaning, it remains an analogical formation.

3.6.5 Reduplication

Reduplication is the repetition of phonological information present in the base lexeme. The reduplicated element can consist of one or more segments, one or more syllables, or the entire string of the base. In addition, it may have its own pre-specified features or segments that interact with the copied material. The reduplicated morpheme can be added as a prefix, suffix, or infix and may serve any inflectional or derivational function that is typical of regular affixation, such as pluralization, distribution, perfective, or progressive aspect, diminution, augmentation, intensification as well as variety and similarity. In (15a), a reduplicated syllable is prefixed to a verb in Mokilese to create a progressive form and in (15b) it is suffixed in Chukchi to signal the absolutive singular. Warlpiri in (15c) reduplicates the entire base in plural formation, while the repeated base in Tamil in (15d) appears with a pre-specified initial segments to signal plurality with variation (Wiltshire and Marantz 2000: 557–61).

(15)	a.	Mokilese:	/wadek/	/wad-wadek/	‘read—is reading’
	b.	Chukchi:	/jil?e-/	/jil?e-jil/	‘gopher, gopher ABS.SG’
	c.	Warlpiri:	/kurdu/	/kurdu-kurdu/	‘child—children’
	d.	Tamil:	/maram/	/maram-kiram/	‘tree—trees and such’

Reduplication was also used in the verb paradigms of Latin, Greek, and Germanic where it formed perfect stems (Latin *curr-/cucurr-* ‘run,’ Greek *lū-/lēlū-* ‘loose,’ Gothic *hait-/haihait-* ‘call’).

In present-day European languages reduplication is mostly found in individual colloquial remnants, cf. English *hush-hush*, *goody-goody*, German *Tamtam* ‘fuss,’ *Pinkepinke* ‘cash,’ French *train-train* ‘routine,’ *trou-trou* ‘row of holes.’ Unproductive rhyme and ablaut variants are also documented: English *willy-nilly*, *rassel-dassle* and *chit-chat*, *flip-flop* and German *schickimicki* ‘fashion buff’ and *Tingeltangel* ‘honky-tonk.’ There are patterns of apparently reduplicative structures that are productive, however, although

their products seem to be predominantly nonce formations rather than suitable candidates for the permanent vocabulary. One type of genuine reduplication is exemplified by creations like English *moon—schmoon*, *baby—schmaby*, *Wittgenstein—Schmittgenstein*, in which the reduplicative copy has a prespecified initial onset *schm-* that replaces the original onset of the base and expresses a pejorative attitude. Another example of this type is the Turkish pattern with the prespecified onset *m-* as in *tabak mabak* ‘plates and the like,’ *dergi mergi* ‘magazines and the like,’ *kapı mapı* ‘doors and the like.’ As for another pattern of apparent whole word reduplication, there is some controversy as to its exact nature—that is, whether it exemplifies reduplication or compounding. While Ghomeshi et al. (2004) refer to it as “contrastive reduplication,” Hohenhaus (2004) sees it as “identical constituent compounding.” In this construction, the head constituent is repeated as a modifier. Much like a regular determinative compound that picks out a subset of the head denotation, an identical constituent compound also identifies a subset of the denotation of its head that contains the prototypical instances of the category. So, one can take a *cat nap* or *sofa nap*, but the identical constituent compound *náp-nap* denotes a real nap, that is, the core sense of the noun. The same goes for *job-job*, *date-date*, and *logic-logic*, cf. also German *Mädchen-Mädchen* ‘real girl (not a tomboy),’ Italian *lana lana* ‘real wool,’ Spanish *mujer mujer* ‘real woman,’ *casa-casa* ‘real house (as opposed to a shelter),’ Russian *zheltyj-zheltyj* ‘real yellow’ and Persian *bikâre bikâr* ‘really unemployed (as opposed to being an artist)’ (Ghomeshi et al. 2004). Ghomeshi et al. show that the modifier constituent within a compound may be inflected, cf. *fans-fans*, *talked-talked*, and that the repetition occurs at the phrasal level as well, cf. *considered-it-considered-it*, *know-him-know-him*. The fact that the process occurs at both the morphological and phrasal levels need not rule out the morphological structures as compounds, however. They have the pre-stress intonation of determinative compounds and express the meaning one would expect if a concept modifies itself, that is the core concept. Furthermore the repeated constituent structure can itself occur as the modifier constituent in a determinative compound, cf. *wórk work day* (Ghomeshi 2004: 333). Expletive insertion (*abso-bloody-lutely*, *Ida-shitty-ho*, *un-fucking-believable*) is often considered a word-formation process expressing emphasis or intensification although it also occurs with the same properties at the phrasal level as in: *That comes as no fucking surprise* or *I’ll bloody swim to Barbados*.

3.7 SUMMARY

The most difficult task in delineating derivation from compounding is in determining when a lexeme has relinquished its independence and become an affix. The discussion of compounding has shown that there is indeed a legitimate intermediate status between a free lexeme and an affix, namely a bound positional variant of a compound constituent. But this is a normal consequence of compounding and should not be misconstrued as affix formation. Other types of bound roots can be found that have arisen due to

individual cases of historical development or to whole-word borrowings from another language. The term affix should be reserved for a bound formative that is not related in a psychological sense with another independent lexeme of the vocabulary. The interesting phenomenon of lexical affixes, that is, affixes that have root-like semantics, does not call this conclusion into question as they are formally speaking undeniably formatives. The number of affixes in a language is finite and their typical phonological and combinatorial properties are known in an intuitive sense to the speakers of the language. This is true also for the incorporating languages that have lexical affixes; their numbers can be large, but they are nevertheless finite. Synthetic compounds are not necessarily ambiguous or hybrid formations, but can be seen as compounds whose head constituents have undergone derivation (often accompanied by argument inheritance) and are subject to normal pragmatic constraints on informativeness. The difference between conversions, back-formations, analogical formations, and blends formed on the basis of a compound and genuine compounds lies in the fact that the meaning of the bi-lexemic derivations is not explainable as a compositional function of the individual constituents, but only by relating them to the whole complex form that serves as their base. Even when a series of like formations occur (cf. *workaholic*, *spendaholic*), their semantic dependency on a base word has a restrictive force on the number of potential formations that are created in opposition to affixation where the affix functions as an independent constituent and is not limited in the same respect. Finally, whereas genuine reduplication patterns with affixal derivation in its formal and semantic properties, many languages allow a superficially similar whole-word repetition process that closely resembles compounding. Rather than expressing a category typical of derivational affixes, the repetition of the head lexeme as its own modifier results in the determinative-like denotation of a core case of the head category.

CHAPTER 4

THEORETICAL APPROACHES TO DERIVATION

ROCHELLE LIEBER

4.1 INTRODUCTION

THE literature on morphology in general and on derivational morphology in particular does not lack for theoretical overviews. Textbooks (Spencer 1991, Booij 2007, Haspelmath and Sims 2010, Lieber 2010a, among others) frequently treat theoretical approaches to derivation thematically, touching on such topics as the nature of word formation rules, level ordering, affix ordering, productivity and blocking, and the like. Several articles in Štekauer and Lieber (2005) give historical overviews of morphological theory (the chapters by Carstairs-McCarthy, Kastovsky, Scalise, and Guevara) or treat particular theoretical models (the chapters by Roeper, Beard, Štekauer, Tuggy, Dressler, and Ackema and Neeleman). My own chapter in Audring and Masini (forthcoming) fulfills this function as well. Therefore in this chapter I will not revisit this familiar ground.

Instead, what I hope to do is to look with fresh eyes at a central theoretical issue that arises especially with respect to derivation as opposed to inflection, compounding, or phrasal syntax. I will frame the discussion in terms of the Saussurean sign, or more accurately in terms of a contemporary re-imagining of the Saussurean sign, as I want to look at both the nature of signifier and signified and the relative importance of the mapping between signifier and signified in the treatment of derivational morphology. In Section 4.2, I look at and recast the Saussurean sign in relation to derivational morphology. In Section 4.3, I introduce the issue of mapping between the signifier and signified of this re-imagined sign. Section 4.4 looks in more detail at the conceptual side of the sign, Section 4.5 at the sensory-motor side of the sign, and Section 4.6 returns to the formal nature of mapping. I will argue that morphological theory has been preoccupied in recent decades with contesting the nature of mapping between signifier and signified, and that when the fundamental computational nature of both the signifier and

the signified is taken into account, the precise formal nature of mapping becomes less important.

4.2 RE-IMAGINING THE SAUSSUREAN SIGN

Most contemporary linguists will be familiar with Saussure's visual representation of the linguistic sign (Figure 4.1).

For Saussure, the fundamental building block of language is the sign, a pairing between a signified or concept, say  and a signifier or sound image, say /kæt/.¹ Saussure himself leaves vague what the conceptual content of the sign is, except to say that it is a segment of thought that is given shape by its pairing with a sound image and that thought itself is "chaotic by nature" (1959: 112).

Re-imagining the basic Saussurean idea in more contemporary terms, we might think of the signifier not so much as a sound image, but as a unit of the sensory-motor system (to use the terminology favored by Chomsky 1995), thus allowing us to speak of language in general, and not just spoken languages (Figure 4.2).

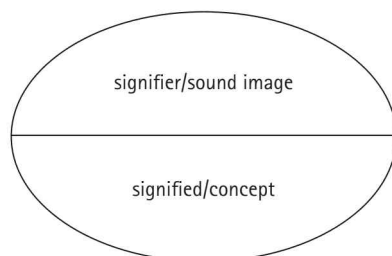


FIGURE 4.1 Saussure's sign

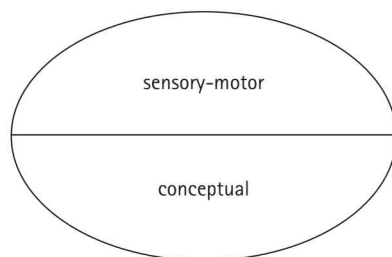


FIGURE 4.2 Saussure's sign re-imagined for the sensory-motor system

¹ Saussure himself uses the image of a tree (1959: 67). I use the cat instead only because it is a symbol that is conveniently available on my computer (and a tree is not!).

It is an uncontroversial tenet of Saussure that the mapping between the signifier and the signified is arbitrary, and this will of course be the case with our neo-Saussurean sign. But even for Saussure, the sign is not always perfectly arbitrary. Saussure had little to say that pertained directly to derivational morphology, but where he comes closest is in his discussion of “relative arbitrariness” (1959: 131–2). To the extent that signs are complex, they exhibit what he calls motivation (1959: 132):

...motivation varies, being always proportional to the ease of syntagmatic analysis and the obviousness of the meaning of the subunits present. Indeed, while some formative elements like *-ier* in *poir-ier* against *ceris-ier*, *pomm-ier*, etc. are obvious, others are vague or meaningless. For example, does the suffix *-ot* really correspond to a meaningful element in French *cachot* ‘dungeon’?

In contemporary terms, we might say that the more segmentable and semantically transparent a complex word, the more motivated or less arbitrary the sign.

Still thinking in more contemporary terms, one central issue that seems to arise in looking at complex words is the nature of non-arbitrariness, or the nature of the mapping between the sensory-motor part of the sign and the conceptual part. Indeed, much of recent morphological theory has been devoted to determining the formal properties of the mapping between the conceptual parts of complex words and the sensory-motor parts. And these formal properties in turn have hinged on the status of the morpheme, specifically whether we conceive of complex words as being composed of morphemes or not.²

4.3 THE PROBLEM OF MAPPING

Morphologists have long been accustomed to thinking of mapping in terms of either rules (analogous to rules of phonology) or hierarchically arranged structures (analogous to syntax). Borrowing from the American Structuralist tradition (Hockett 1954), these have been referred to respectively as Item and Process theories (IP) and Item and Arrangement (IA) theories. However, a useful and somewhat more sophisticated conceptualization of different models of mapping is that of Stump (2001), who has proposed a taxonomy based on two cross-cutting characteristics.

According to Stump, morphological theories can first of all be characterized as either LEXICAL or INFERENTIAL. In a lexical system, Stump characterizes “the association

² Saussure does not use the term “morpheme” and his position on the status of complex words is equivocal, as Carstairs-McCarthy (2005: 7–9) has shown. On the one hand, there are parts of the *Cours* where it appears that Saussure works with something like a notion of the morpheme, so that a complex word like *happiness* would consist of two separate signs in a structural or “syntagmatic” relationship to one another. On the other, it more often seems that Saussure treats complex words as whole signs, with their internal structure emerging as a function of what he calls “associative” relations to other signs (for example, *redness*, *baldness*, *hardness*, *squareness*, ...), as opposed to syntagmatic relationships.

between an inflectional marking and the set of morphosyntactic properties which it represents as being very much like the association between a lexeme's root and its grammatical and semantic properties." In an inferential system, "the systematic formal relations between a lexeme's root and the fully inflected word forms constituting its paradigm are expressed by rules or formulas" (2001: 1). Lexical theories embrace the morpheme as a unit of structure and inferential models do not. A second, and orthogonal, dimension of Stump's taxonomy divides mapping systems into those that are INCREMENTAL in the sense that "words acquire morphosyntactic properties only as a concomitant of acquiring the inflectional exponents of those properties," and those which are REALIZATIONAL, where "a word's association with a particular set of morphosyntactic properties licenses the introduction of those properties' inflectional exponents" (2001: 2).

Stump argues that *inflectional* morphology is best served by a model that is inferential and realizational, for example, his own Paradigm Function model. In contrast, theories that build inflected forms from inflectional morphemes that are put together with bases via syntactic or quasi-syntactic rules might be characterized as lexical-incremental models; Lieber (1992) is a model that takes this form. Distributed Morphology (Halle and Marantz 1993) represents a combination of lexical and realizational features, and Steele's (1995) "Articulated Morphology" is inferential and incremental, according to Stump (2001: 1–3). At the forefront of all of these models is the precise formal characterization of a mapping between form and meaning, viewed primarily from the perspective of inflection.

Stump's model is an excellent one for considering the nature of mapping in inflection at least in part because in the case of inflection we have a reasonably good

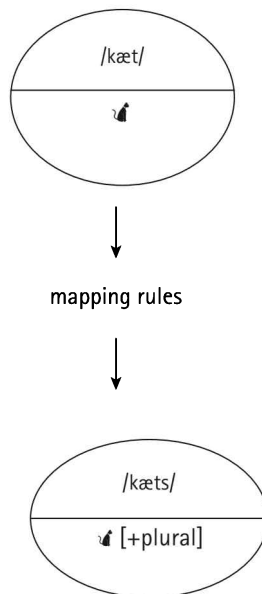


FIGURE 4.3 Mapping in inflection

characterization of what we are mapping onto what. That is, confining ourselves for the sake of convenience to spoken language, it is relatively uncontroversial that for inflection the mapping system must pair morphosyntactic features with phonological forms. Further, we have a pretty good idea of what morphosyntactic features look like. We may quibble about how many features are necessary and what their values should be, but it is not controversial to assume that there are number features, person features, tense features, and the like. We might visualize the theoretical treatment of inflection as in Figure 4.3.

We can, of course, look at the formal nature of *derivation* as well in terms of Stump's lexical/inferential and realizational/incremental parameters. But with regard to derivation, we do not have nearly as clear a notion as we do with inflection what we are mapping onto what. We will focus in the next section on the conceptual side of the sign and return to the formal complexity of the sensory-motor representation in Section 4.5.

4.4 THE CONCEPTUAL SIDE OF THE SIGN

The nature of what we might call the derivational signified has typically been left vague in most formal treatments of derivation. Indeed, as I argued in Lieber (2004), there is little agreement in the literature on what the semantic representation of simplex words should be, much less how the semantic representation of simplex words compares to that of derived words. Theorists like Anderson (1992) and Stump (2001) have assumed that derivation would be well-served by the realizational/inferential model. This conclusion is justified to the extent that the conceptual side of the derived sign is analogous to that of the inflected sign. But is it?

The classic Saussurean diagram implies that the conceptual part of the sign is in some way a holistic image that we represented earlier as . Saussure seems to claim that thought is chaotic and is organized only in the pairing of the signified with the signifier. One particularly contentious matter in the interpretation of Saussure is whether the signified is to be identified only by virtue of its relation to other signifieds—Saussure's notion of “value”—or whether there is something positive to be attributed to the signified as well. Although many interpretations of Saussure privilege the notion of “value,” I take the position of Bredin (1984: 72) that there must be more to the signified than what it is not. As Bredin argues, “When it is said that a concept is defined by its ‘not being’ any other concept, this is a shorthand, perhaps misleading, way of saying that it occupies a different place in the language system from any other.” The point is that whatever (or wherever) that “place” is, it is something positive.

In my re-imagining of the sign, I will concentrate on the positive content of signs as opposed to the “value” that arises from their relationship to one another. The primary question we must raise is how the conceptual content of the neo-Saussurean sign should be characterized formally. There are many conceivable formal models, including model theoretic semantics, the Lexical Conceptual Structures of Jackendoff (1990

and subsequent work), or the Natural Semantic Metalanguage of Wierzbicka (1996), but here I will fall back for the sake of illustration on my own framework (Lieber 2004, 2006, 2009a, b).

The symbol ♣ suggests that there are certain aspects of our knowledge of the concept “cat” that are sensory in nature—visual, tactile, auditory, and so on. This is what has been called in the literature “encyclopedic” knowledge (Harley and Noyer 1999, 2000), the “constant” (Rappaport Hovav and Levin 1996, 1998), “Conceptual Structure” (Mohan and Mohan 1999) or the “semantic body” in my own work. But as I have argued at length in Lieber (2004) and elsewhere, this encyclopedic knowledge of lexical meaning is only one part of our knowledge of lexical semantic representation. It is relatively uncontroversial that there is also a more formal and conventionalized part of meaning that we need to attend to, what I have termed the “semantic skeleton.” In the framework of Lieber (2004) the skeleton conveys the meaning at least that “cat” is an item that is referential in nature, and moreover one that is concrete and not processual. Theorists might disagree on the nature of the primitives that constitute the more formal part of the semantic representation, as well as on the way they are analyzed or generated, but I argue in Lieber (2004) that any account of the conceptual side of the sign must have something to say about both formal and encyclopedic aspects of meaning.

Another point that is critical for our purposes is that the semantic representations of underived words are not necessarily simple, but may in fact be built up out of smaller primitive parts. Further, those parts do not occur as an unstructured mass, but must be ordered in some way: linearly, hierarchically, or both. In other words, there must be some sort of structure to conceptual representations. And if this is the case with underived signs, it must also be the case—surely even more so—with derived signs. Further, assuming that parsimony is to be desired in a morphological theory, whatever the system for constructing the conceptual representations for underived signs turns out to be, the same sort of conceptual representations ought to be useful for derived words as well.

What follows is a short sketch introducing the framework of lexical semantic representation elaborated in Lieber (2004, 2006, 2009a, b, 2010b). In Lieber (2004: chs 1, 2) I argue that any framework for the representation of the semantics of words (what I have been calling here the conceptual part of the sign) must have several features:

- it must be decompositional
- its primitives must be of the right “grain size”
- it must be cross-categorical
- it must be able to deal with both simple and complex lexemes

To these desiderata I would now add another. Although it is implicit in the idea of decomposition that what can be decomposed must have been composed in the first place, let me make explicit here that a theory of lexical semantic representations must have some sort of rules for composing primitives into well-formed representations. Given the possibility of creating potentially infinite numbers of newly derived words,

we must assume that there is a computational aspect to derivation and that some of that computation is semantic.

Suppose then that our framework contains some system for generating conceptual or semantic representations, which I will henceforth refer to as skeletons. Following most systems of this sort (model theoretic semantics, Jackendoff's Parallel Architecture 1990, 1997, the system of Lieber 2004), we will assume that skeletons consist of functions that take arguments. Abstracting for the moment away from the precise nature of semantic functions, we will assume first that our system generates representations of the following sort:

- (1) SKELETON $\rightarrow$ [F (arg)]
 SKELETON $\rightarrow$ [F (arg, arg)]
 SKELETON $\rightarrow$ [F (arg, arg, arg)]

That is, functions can take up to three arguments. Why three is the upper limit is a question we might wonder about, but observation of semantic representations in the literature suggests that three arguments will be sufficient for our purposes.

Arguments, in turn, can be either open slots in the representation, which we will represent with square brackets, or they may themselves be skeletons. In other words, skeletons are recursive:

- (2) arg $\rightarrow$ []
 arg $\rightarrow$ SKELETON

Open slots will be satisfied in various ways, for example by being linked to a syntactic phrase or coindexed with another open slot in a skeleton, as will be illustrated in (4) and (5) below.

Functions would ideally be limited to a finite number of primitives. In Lieber (2004) I sketch a highly constrained featural system that is appropriate for lexical (as opposed to grammatical) meanings. Here, we need not worry about the nature of those features, although we will return to them briefly in Section 4.6. For now we will represent functions schematically with Greek letters:

- (3) F $\rightarrow \alpha, \beta, \gamma, \delta, \dots$

The final part of the system that is necessary, at least within the framework of Lieber (2004), is a means of integrating the skeletons that are composed as part of the word formation process. Within the literature on word formation, this sort of integratory principle is usually represented as some sort of coindexation, roughly speaking, a process that identifies arguments in a skeleton as being matched with the same referent. Exactly how coindexation works is an important issue, but for our purposes a more or less generic version such as that in (4) will be sufficient.

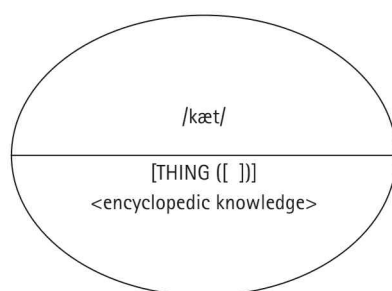


FIGURE 4.4 The simplex sign *cat*

(4) Principle of Coindexation

In a configuration in which semantic skeletons are composed, coindex the highest argument with the highest (preferably unindexed) embedded argument. Indexing must be consistent with semantic conditions on arguments, if any.

The rules in (1)–(4) would then give us well-formed skeletons like those in (5):

- (5) a. $[\alpha ([])]$
 b. $[\alpha ([_i], [\beta ([_i], [])])]$

Assuming, then, that skeletons like those in (5) are on the right track, and are associated as well with encyclopedic knowledge, then a simplex sign like *cat* would look something like that in Figure 4.4, glossing again over the nature of the primitives that constitute functions and representing them as in a familiar sort of shorthand as *THING*, *CAUSE*, *BECOME*, *STATE*, and so on.³

In other words, the sensory-motor part of the sign */kæt/* is associated with the conceptual part of the sign, or skeleton, in an arbitrary way.

Of course, the skeleton need not itself be as simple as that in Figure 4.4. Suppose that we look instead at an underived sign like *kill* (Figure 4.5).

Although the sensory-motor side of the sign is comparable to that in Figure 4.4, because *kill* is a verb and specifically a causative verb, the conceptual side is considerably more complex, with function embedded within function embedded within function (see Lieber 2004 for a full treatment of causative verbs). It is important to note here that causative verbs are semantically complex whether or not they are morphologically complex. That semantic complexity is ultimately the result of applying rules like those above for composing skeletons.

³ In the theory of Lieber (2004), what we are representing here as *THING* would be represented as the semantic feature [+material], *CAUSE* would (glossing over some details) be [+dynamic], *BECOME* would be [+dynamic, +IEPS], and *STATE* [–dynamic]. Again the precise nature of the featural system is not important to the issue at hand.

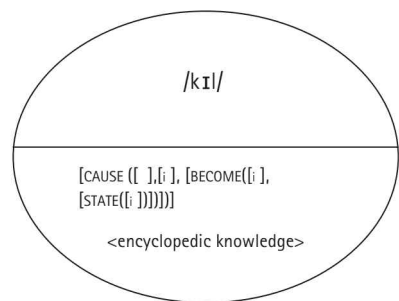


FIGURE 4.5 The simplex sign *kill*

We now have an interesting dilemma on our hands, that is, the dilemma of what constitutes a complex sign. Traditionally we have thought of complexity primarily as a matter of morphological segmentation (that is, a complex word has more than one morpheme). But it seems that there is more to complexity than this. Indeed, the notion of complexity is highly theory dependent. We must acknowledge that the conceptual side of the sign can be complex in its own way. Let us define a complex skeleton as one in which a function is embedded under another function. A simplex skeleton is one that contains only a single function with its (non-function) argument(s).

In addition to semantic complexity, we must also recognize sensory-motor complexity, a point we return to in more detail in Section 4.5. For example, /kɪl/ consists of a single morpheme (indeed, of a single syllable), arguably a sensory-motor representation that is relatively simple. But the sensory-motor representation can of course be complex as well, and indeed simple sensory-motor representations can be mapped onto either simple or complex skeletons, and complex sensory-motor representations may be mapped onto skeletons that are either simple or complex. As illustrated in (6) we have four possibilities:

(6)	sensory-motor		conceptual
a.	simple	↔	simple
b.	simple	↔	complex
c.	complex	↔	complex
d.	complex	↔	simple

We have considered cases (6a) represented by *cat* and (6b) represented by *kill*. We will now consider several examples of (6c), before considering whether (6d) is a plausible scenario.

In order to consider all the various possibilities, let us look in detail at two words that share core components of meaning with the simple sign *kill*, namely *deadify* and *euthanize*. The form *deadify* is a neologism listed in *Urban Dictionary* (<<http://www.urbandictionary.com/>>) with the meaning ‘to make someone dead, to kill, to own noobs’; as of

January, 2013, it is not attested in the *OED*.⁴ The words *kill*, *deadify*, and *euthanize* differ, of course, in the encyclopedic aspects of their meanings: *kill* is relatively neutral, but *euthanize* involves killing a person or animal to put an end to their suffering, and *deadify* seems to be associated in the examples in *Urban Dictionary* with the sort of killing that goes on in video games, “noobs” referring to inexperienced players. Both *deadify* and *euthanize* may be considered as hyponyms of *kill*, but I will assume that the hierarchical relationship is a function of the semantic body rather than the semantic skeleton.

What is interesting for our purposes is that these words differ in the encyclopedic components of their skeletons, and also arguably in the complexity of their morphological structure, that is, in the sensory-motor component of our re-imagined sign, but not in the complexity of the skeletal part of their skeleton. *Deadify* is nicely compositional, with both base and affix having easily identifiable forms. In a lexical theory that countenances morphemes and a complex syntactic structure of words we might represent the semantic composition of those morphemes as in Figure 4.6a.

Assuming such a morphemic analysis, we might imagine that the rules for composing skeletons embed the skeleton of *dead* in that of *-ify*, and indexation occurs to integrate the two representations into one. The composed skeleton is precisely the same as that for *kill* in Figure 4.5.

This analysis is not the only possible one for the sensory-motor side of the sign. Within inferential models, there is no reason to demand that the sensory-motor part of the sign be seen as structurally complex or be composed by syntax-like rules Figure 4.6b.

The composition of the skeleton is precisely the same, but in an inferential model, we make no claims as to any internal complexity for the phonological form of *deadify*. The complex skeleton is mapped onto a word that lacks internal morphological structure.

The word *euthanize* can be treated in a similar fashion. In this case, the affix is clear in form, but the base is not free-standing in English, and is therefore of dubious status, even from the point of view of lexical models. But again the internal composition of the sensory-motor form *euthanize* need not trouble us. Quite apart from whether we think that the word should be parsed as *euthan* + *ize* or not, the skeleton of the complex word must have a composed form that is precisely the same as that for *deadify* and indeed for *kill*. That is, although the sensory-motor part of the sign may be simple or not, the conceptual side of the sign is complex, just as it is in *kill* or *deadify*. We can represent this either as Figure 4.7a or as Figure 4.7b, where the only difference with the skeleton is whether the pieces are associated with the word as a whole or with two separate sensory-motor representations. Either way, the parts of the skeleton are composed by embedding [STATE ([])] in open slot labeled <base> and coindexing arguments.

We have now looked at three of the four cases in (6), the cases in which a simple sensory-motor representation is mapped onto a simple skeleton or a complex skeleton, and two types of cases in which a complex (or potentially complex) sensory-motor representation is mapped onto a complex skeleton. To be thorough, we should consider the

⁴ There is one attestation in COCA, but it seems fairly clearly to be a typographical error.

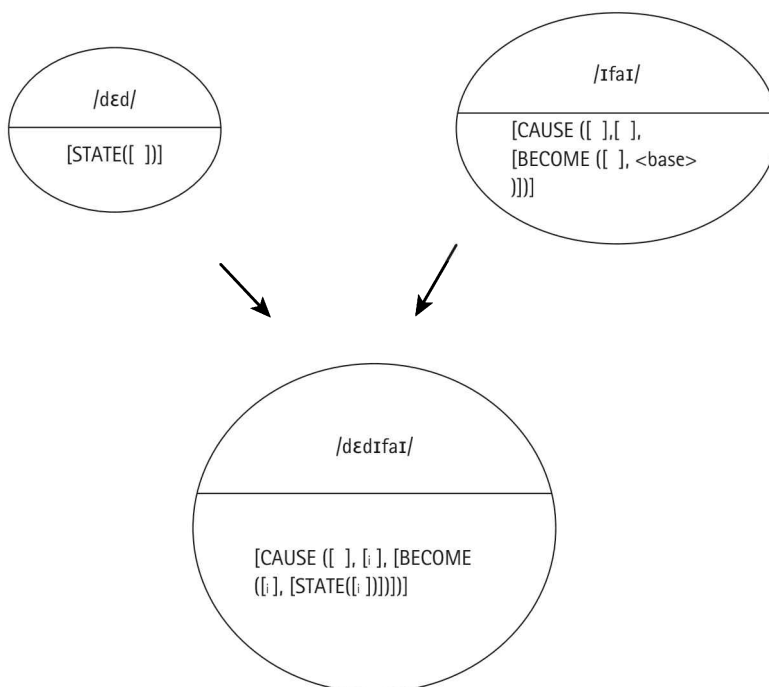


Figure 4.6a Semantic composition in a lexical model

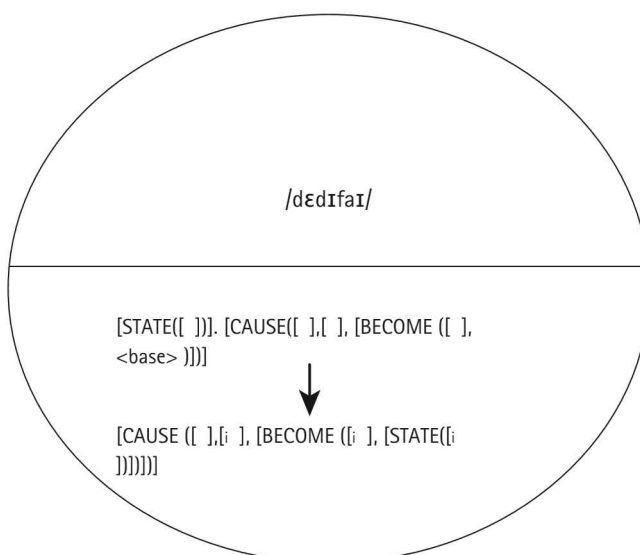


Figure 4.6b Semantic composition in an inferential model

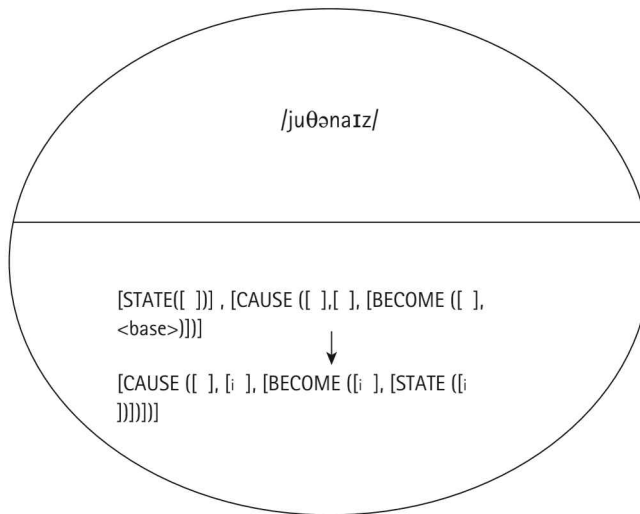


Figure 4.7a Euthanize in an inferential model

possibility of the fourth case, where a (potentially) complex sensory-motor representation is mapped onto a simple skeleton. This is what we would get in the case of a derived word with a highly lexicalized meaning, as for example in the case of the word *transmission* with the meaning ‘gearbox.’ Here, the sensory-motor part of the sign might plausibly be analyzed as [[transmit]ion], but the skeleton for the word would consist (at least in the system of Lieber 2004), as a single function with a single argument: [THING ([])].

4.5 THE COMPLEXITY OF THE SIGNIFIER

We have thus avoided discussing the form of sensory-motor representations, assuming them to be better understood than skeletons. It would, of course, be disingenuous to believe that the nature of sensory-motor representations is any less complex or interesting than that of the conceptual side of the sign. Although it is beyond the scope of this chapter to work out details of a theory of morphophonological representation, it is nevertheless worthwhile to consider at least briefly the nature of the complexity we find in the sensory-motor part of our sign. Presumably most linguists would agree that the smallest units of phonological structure are distinctive features, which are organized into segments, which in turn are organized into higher prosodic units such as syllables, feet, and prosodic words. Further, we must take into account that the phonological form of complex words is not necessarily concatenative, so that other factors such as reduplication, input and output conditions and templates are at play in the derivation of complex words involving reduplication or subtractive processes, for example (see Plag 1999, Lappe 2003, Inkelas and Zoll 2005, McCarthy 2008, as well as chapters 11 and 12 in the present volume). Confining ourselves henceforth to concatenative morphology, we might articulate the sensory-motor part of a complex word

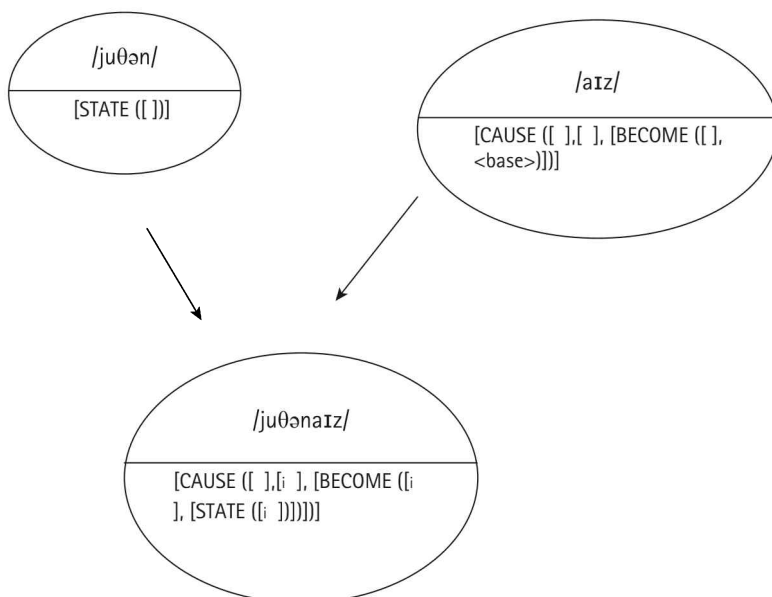
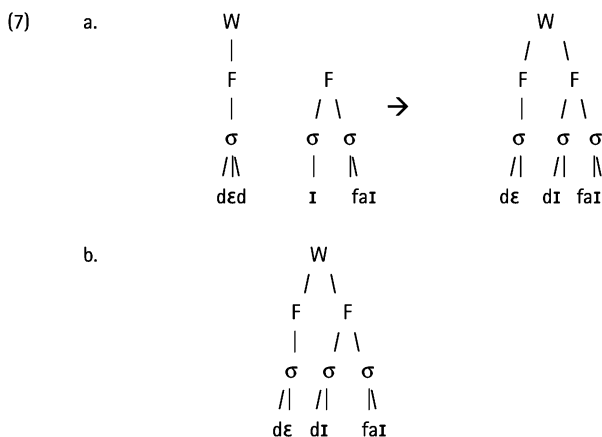


Figure 4.7b Euthanize in a lexical model

more fully as either (7a) or (7b), depending on whether we follow a lexical or an inferential conception of morphology:⁵



The representation in (7a) is, then, an elaboration of the sensory-motor representation in Figure 4.6a, and (7b) of that in Figure 4.6b. What is important to point out in either case, is that the phonological structure is not necessarily isomorphic with the conceptual structure of the sign, as segments may, for example, syllabify across morpheme boundaries.

⁵ For reasons of space, we start at the level of the segment in these representations, rather than with distinctive features.

The observation that complex words may have more than one sort of structure and that these structures need not be isomorphic has been made many times in the literature, figuring prominently in the debate over bracketing paradoxes (see Booij and Lieber 1993). We should observe, though, that the mismatch in structure has traditionally been stated in relatively simplistic terms. For example, the paradox concerning the word *unhappier* is frequently observed to be that the semantic bracketing must be [[un[happy]]er], whereas the phonological bracketing must be [un[[happy]er]] (see for example Spencer 1991: 44). But of course each of these structures can only be understood as shorthand for more complex and articulated conceptual and sensory-motor structures, perhaps of the sorts I have discussed here.

What I have tried to do so far is to refocus our discussion so that we continue to think about mapping, but also to take into account the ways in which both the sensory-motor part of the sign and the conceptual part of the sign may be complex, as well as the fact that complexity on one side need not be matched with complexity on the other. In this light, it then makes sense to return to the issue of mapping.

4.6 A RETURN TO MAPPING

Thus far I have argued that both the conceptual and sensory-motor parts of the sign must be re-imagined as highly structured and that we are now approaching the point where we can consider how the mapping between these highly structured entities is to be accomplished. Before we do so, however, there is still one step that I think we have missed, namely that we must think about what the mapping process needs to accomplish. Specifically, what information does mapping add that cannot be gleaned either from the highly articulated phonological representation or from the skeleton?

Consider the sort of information that is available in what is typically called the “subcategorization frame” in a lexical model of derivation. I will use as my example the sort of word-syntactic theory I developed in Lieber (1980, 1992) in which morphemes are assumed to have lexical entries of the sort in (8) (Lieber 1980, 66):

- (8) dead [A]
 (phonological representation)
 semantic representation: ...
 -ify]_{N,A} —]_V
 (phonological representation)
 semantic representation: causative

The subcategorization frames in (8) tell us three things: whether a morpheme is free or bound, if bound, whether the morpheme precedes or follows its base, and finally, the categorial identity of the base and the resulting derived word. Almost the same information

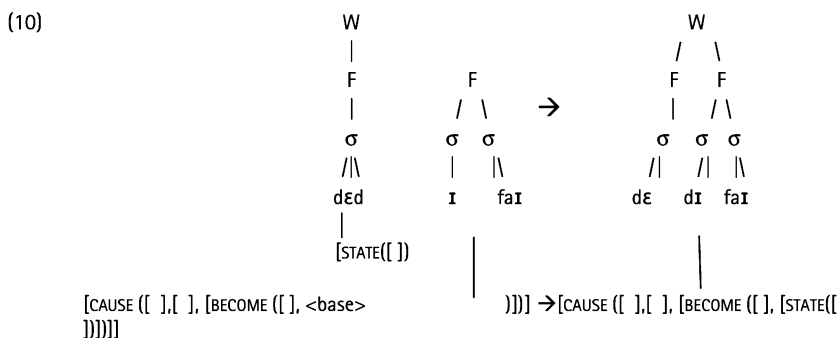
can be gleaned from the representations of a realizational model, as (9) illustrates (from Stump 2001: 257):

- (9) a. PF (<friend, privative adjective>) = <friendless, privative adjective>
 b. PF (<X, privative adjective>) = <Xless, privative adjective>

PF stands for “paradigm function” in Stump’s model, the equivalent of a derivational rule. As in the subcategorization frames of the lexical model, linear order is made explicit by the rule. Note that although Stump does not give the categorial identity of X, *Xless* is specified as an adjective, and specifically a privative adjective, combining categorial and semantic information. The realizational model differs from the lexical in not explicitly identifying *-less* as a bound morpheme, as this model does not make use of the notion of “morpheme” as a primitive, but the boundness of *-less* can be inferred from its presence only on the right side of the equals sign. Note as well that in both the lexical and the realizational type rules, neither the phonological nor the conceptual representations are formalized, so their complexity is left implicit.

The question we now return to is this: given a sufficiently articulated formal representation of the sensory-motor and conceptual representations, can any of this “mapping” information be inferred from what we already have? To the extent that some of this information follows automatically from other aspects of our representations, we will be able to reduce the importance of mapping issues. Further, if mapping cannot be completely dispensed with, we are left with the question of what is the best way to represent the mapping residue. This is the issue to which we now turn.

If we consider again the nature of phonological representations and skeletons, it appears that linear order can be inferred from the phonological representation of the sign, as Sproat (1985: 78) has already suggested. Further, the theory of lexical semantic representation that I have sketched in Section 4.4 assumes that hierarchical structure is encoded into the structure of the skeleton.



If phonological representations encode linear relations and skeletons encode hierarchical relations, it is conceivable that all that is left of the traditional subcategorization

frame or paradigm function rule is information about categorial selection. Where should that be encoded?

In Lieber (2006) I tried to argue that at least some categorial information can be derived from skeletons. All skeletons whose outermost function contains the feature [material] will correspond to nouns, and all skeletons whose outermost function consists of the feature [dynamic] without [material] will correspond to verbs or adjectives. Skeletons that lack both [material] and [dynamic] (but are characterized by other features) will correspond to adpositions. To the extent that such a program is successful, we might succeed in removing the last part of the mapping residue, with the result that the mapping between our re-imagined signifier and signified is unmediated by any other structure. That would be a startling conclusion, of course, as it would imply that no separate mapping rule or process is needed at all. But this is no doubt too strong a conclusion to draw.

Linguists have long attempted to derive syntactic categories from notional or semantic categories, but such attempts have never been entirely successful. Not surprisingly, the program of Lieber (2006) faces problems, not the least of which is its difficulty in distinguishing stative verbs from adjectives in any straightforward way (both are characterized by the feature [-dynamic]). If syntactic categories cannot in the end be derived fully from semantic categories, the mapping problem will never disappear entirely.

But suppose that it does not. It will still be greatly diminished. If both linear order and hierarchical structure are derived from other necessary parts of the representation, what we are left with is the task of providing the categorial information that mediates between the two sides of the base sign and those of the derived word (in an inferential framework that does not recognize morphemes) or between the two sides of the base and affix signs (in a lexical framework that does). We can look at complex words either way, but it is difficult to see what the empirical differences between the two approaches might be.⁶

4.7 CONCLUSION

In this chapter I have not provided a conventional historical survey of theoretical developments in morphology as they concern derivation. Rather, in reimagining the Saussurean sign in contemporary terms and applying it to complex derived words, I have tried to take a broad meta-theoretical approach; my hope has been to re-examine a specific preoccupation of our field over the last three decades. I have argued that we have tended to concentrate our attention on the issue of mapping without first adequately exploring what we are mapping. Theorists have especially given short shrift to the nature of the lexical semantic representation. What I have tried to establish is that

⁶ A similar point is made in Sproat (2005).

the theory of derivation must attend to the complexity of both the sensory-motor and conceptual sides of our re-imagined sign, and that by doing so we might go beyond the issue of the formal nature of mapping between the signifier and the signified that has so preoccupied morphologists over the last three decades. It remains to be seen where this redirection might take us.

CHAPTER 5

PRODUCTIVITY, BLOCKING, AND LEXICALIZATION

MARK ARONOFF AND MARK LINDSAY

5.1 INTRODUCTION

THE topic of morphological productivity as it has been conceived in linguistics for the last half-century is treated in greatest detail in Bauer (2001). If our brief discussion here leads the reader to that book, we will have gone a long way to doing our job. In this chapter, though, we also have a different aim, which is to recast the problem of morphological productivity in a different light. Indeed, we aim to show that the term itself may sometimes be less than helpful. We believe that the most interesting and, more importantly, addressable questions in this domain have always involved not the somewhat elusive notion of productivity, but rather competition. Before getting to that point, however, we must address a more fundamental question, one whose conventional response in linguistics has impeded progress in this particular domain, although it has been enormously helpful in the investigation of other areas of language, whether linguistic systems are entirely discrete in nature.

5.2 IS LANGUAGE DISCRETE?

The success of modern linguistics has always been rooted in the realization that languages are systems. But what sort of systems? Linguistics historically has been most successful dealing with discrete patterns and so we tend to assume that languages are wholly discrete systems. The analysis of productivity in word formation presents one of the most serious challenges to date to the blanket assertion that all patterns in language are discrete.

The first great achievement of historical linguistics in the 19th century, the comparative establishment of the relationships among the members of the Indo-European and Uralic language families, was based on the discreteness or “regularity” of sound change, what were called phonetic laws and compared at the time to the material laws of science. Famously, the major exceptions to Grimm’s law in Germanic languages were later subsumed under Verner’s law, providing strong confirmation for the methodological assumption of the regularity, exceptionlessness, or discreteness of sound laws.

In the first half of the 20th century, regularity and discreteness again showed great success, this time in the analysis of phonological patterns, the discovery of the phoneme, and the categorical nature of phonological alternations. Phonemic contrasts are famously categorical and even the distribution of allophones is usually taken to be discrete.

The second half of the last century saw the ascendance of syntax. The immediate constituent analysis of Rulon Wells (1947) led quickly to the phrase structure grammar of Chomsky (1957), which formed the foundation for transformational grammar, all discrete systems. All prominent frameworks for syntactic analysis since then have been discrete.

By the 1960s, especially with Chomsky’s (1965) distinction between competence and performance, most linguists could presume comfortably that language was rule governed at its core, so that all components of grammar could be assumed in turn to be discrete systems of regular rules. The messy nondiscrete aspects of language could be relegated to matters of performance or the lexicon, which Bloomfield (1933) had already characterized as a list of irregular items.

Morphology is a challenge for any theory of language that is focused on discreteness and regularity, because so much of morphology is neither. The first challenge for morphologists is to figure out how to integrate regular and irregular phenomena. In inflection, the tried and true method of assuming the dominance of regularity that had succeeded since the days of the Neo-grammarians again proved successful. A variety of researchers from Aronoff (1976) to Pinker (1999) to Brown and Hippiusley (2012) worked out the idea that irregular items, listed in the lexicon, could preempt or *block* their regular counterparts, which would emerge as defaults when not preempted by the irregulars. So, the English irregular past tense form *sang* blocks the regular form **singed*, which is the product of the default rule for past tense that adds the suffix *-ed* to English verbs.¹ There is even a hierarchy of outright exceptions like *went* instead of **goed*, rules with narrowly specified domains like the ablaut rules that characterize the relations among *sing*, *sang*, and *sung*, and the default regular rules.

One way to look at these narrow domains is in terms of the scope of the rules or relations that characterize them. A form like *went* is not describable by any synchronic

¹ In actuality, matters are not so simple. Bauer et al. (2013) show that many English irregular verbs show variation between regular and irregular forms, some well-known cases being *dived* vs. *dove*, *lighted* vs. *lit*, and *shined* vs. *shone*. For all these, variation is documented from quite early.

generalization that goes beyond one verb; *went* must simply be listed as the past tense of *go*.² A form like *transcended*, by contrast, is most easily thought of as being rule-derived, in the same way as a sentence like this one must be rule-derived. But the few hundred irregular verbs of English can be thought of as either stored or, more palatably for some, characterized by rules, if we assume that these rules simply have narrower scope than the default rule. The value of the rules for the linguist is that they express the generalizations, admittedly limited, that can be extracted from this set of irregular verbs. The potential wider applicability of these rules is revealed in the errors of children and second-language learners, who may produce forms like *brang*, extending the rule for *sing* to *bring*.

If these irregular but not completely unpredictable phenomena could truly be cast purely in terms of increasingly larger domains, then we could call productivity a discrete phenomenon and preserve the claim that the entire core of language, linguistic competence, is discrete. Unfortunately, the tactic fails. It is not just that the set of *sing/sang* verbs is limited to monosyllables ending in the sequences *-ing* and *-ink*. More importantly, not all such monosyllabic verbs succumb to the rule. Consider, for example, the three homophonous verbs *ring* (*my bell*), *ring* (*the city*), and *wring* (*out the clothes*). Each has its own distinct past tense forms: *rang*, *ringed*, and *wrung*. We might be able to tag *ringed* as an exception to the smaller-domain rule, so that it then falls under that larger-domain default, but we cannot do that with *wrung*, which must be either lexically listed like *went* or marked as exceptionally showing the vowel that we find in *hang/hung* instead of the vowel that is “normal” in irregular verbs ending in *-ing* and *-ink* like *sing* and *sank*. Furthermore, new verbs of this form are invariably subject to the default rule: *clinked*, *dinged*, and the website *Blinged out Blondes*, from which all things rhinestone are readily available.

Our inability to cast these phenomena in terms of domains leads next to the notion of discrete degrees of productivity. Default rules like *-ed* are *fully productive*: they apply to any verb that they encounter, except for those that are covered by rules of narrower scope. The rules for the past tense of English strong verbs, by contrast, are less than fully productive: they do not apply to every verb that meets their conditions. We can call them *semi-productive*. But now we need to ask ourselves how many of these discrete degrees there are. As Bauer (2001) so eloquently shows in his chapter on degrees of productivity, this question of how many degrees there are leads to a slippery slope that results inevitably in the abandonment of discreteness as a solution to the problem of productivity in word formation.³ Bauer’s catalog of terms that linguists have used for intermediate degrees of productivity includes, besides *semi-productive*, *semi-active*, *active* (though not fully productive), and *marginally productive*. We are led in the end to conclude, as Bauer does, that morphological productivity is scalar rather than discrete and that there is no finite number of degrees of productivity for us to name. As with the points on a compass, we may begin by naming four directions (North, East, South, and West) then

² The verb *go* has lacked a morphologically related past tense form since earliest Germanic times, at least. The past tense form *went* is from the verb *wend*, which is uncommon in Modern English but has a regular past tense form *wended*.

³ The claim that productivity is categorical is asserted as recently as Yang (2005).

add the four intermediate points Northeast, Southwest, Northwest, and Southeast, but soon we find ourselves needing to talk about South Southwest. Eventually we divide the circle of the compass into 360 degrees, each of these is divided into minutes and then seconds, but in the end we give up and admit that the points on the compass, as with any other circle, are numberless.

5.3 SCALAR PRODUCTIVITY VERSUS BLOCKING IN WORD FORMATION

The domain of linguistic inquiry in which the scalar nature of morphological productivity emerges most clearly is that of word formation or lexeme formation (Aronoff 1976, 1994). This was the last core aspect of language to be investigated by modern theoretical linguistics, most likely because of its resistance to discrete methods of study. The treatment of the relative productivity of rival English suffix pairs in Aronoff (1976) provides a valuable history lesson. The unconscious assumption underlying the entire discussion is that the difference between the two rivals is scalar rather than discrete, but the intellectual climate of the time made it impossible for this assumption to be made explicit even to the author of the work, as one of us can attest personally. It was only some years later that the author could even begin to formulate it (Aronoff 1983).

Aronoff (1976: 43) attempted to reduce the contest between rival suffixes to what was termed *blocking*, defined there as “the non-occurrence of one form due to the simple existence of another,” a definition since subject to much discussion and some revision (Rainer 1988, Bauer 2001). Blocking, understood in that sense, is discrete: one form exists and the other does not. But later research, some of which we discuss in more detail below, has revealed that this discrete definition fails to capture most of the more subtle interactions that we would surely like to subsume under the term. Most notably, as van Marle (1985) and Rainer (1988) observe, we would like to account for the rivalry within pairs (or larger sets) of affixes, not just between pairs of words, as this definition does. Furthermore, when one word blocks another, the blocked word may still occur, sometimes not with the sense that would be assigned to it if it had no rival, again contrary to this simple definition. In a sense, the word may be deflected instead of blocked.

Here are a few simple examples of how a rival word may be deflected rather than simply blocked. Consider the three English affixes *-ness*, *-ce*, and *-cy*. We can see from the three words *pleasantness*, *elegance*, and *buoyancy* that they can be rivals, each forming abstract nouns from adjectives.⁴ We know that *-ness* is overwhelmingly the overall default suffix for forming abstract nouns from the entire domain

⁴ There are others, most notably *-ity*, whose competition with *-ness* is the standard example. But *-ity*, like the two other Latinate suffixes mentioned here, is morphologically conditioned. It does not attach to

of adjectives but each of the other two suffixes can be more productive than *-ness* in restricted domains. While *-cy* is the least common overall, it is the most favored of the three with the few words ending in *-ate*: *piracy* (**pirace*), *profligacy* (**profligace*), *delegacy* (**delegace*). By contrast, *-ce* is the most productive with words ending in *-ent* and *-ant*: *diligence*, *dependence*, *resistance*. But in neither of these cases can we say that the rival suffixes are always completely blocked. Sometimes, the *-ncy* rival word is more acceptable than its *-nce* counterpart: *incumbency* (with about 134,000 Google hits) vs. ?*incumbence* (with only 230 or so Google hits). The *OED* lists a fair number of *-nce/-ncy* pairs, and asserts that the former expresses more distinctly the sense of action or process, while the latter expresses the sense of quality, state, or condition, citing the pairs *coherence/coherency*, *persistence/persistency*, and *compliance/compliancy*. What we actually find is no overall generalization but rather that, when both members of any given pair are entrenched, there is often a difference in meaning and the overall less productive *-ncy* member of the pair conveys a more specialized sense. Compare *excellence* with *excellency*. Both words have a long history in English but *excellency* has come to be used largely in honorific expressions like *your excellency*.⁵ *Excellentness* is listed in the *OED* as obsolete, though it shows about 10,000 hits on Google, about a quarter of them from fans of *Bill and Ted's Excellent Adventure*, the classic 1989 cult movie. The pair *compliance* and *compliance*, cited as an illustration in the *OED*, with *compliance* supposed to signify the action or process and *compliance* the quality, state, or condition, have both become much more popular in the last century than they were in the 19th because of the importance of the problem in modern bureaucracies. We do not find, however, that the *OED* distinction holds at all in real examples. A cursory examination of actual Google citations reveals that *compliance* has a more technical flavor and is used for foregrounding and naming: "Let the Compliance Group solve your compliance puzzle." (<compliance-group.com>). This is what one would expect from the fact that *-ncy* is overall the less productive of the two suffixes (Aronoff 1983). The same is true for *dependence* and *dependency*. Both have *OED* citations dating to the 16th century and many of their senses have overlapped since then. The latter, however, most frequently signifies "A dependent or subordinate place or territory; *esp.* a country or province subject to the control of another of which it does not form an integral part" (*OED* online). This is a highly specialized concrete sense very far from the general abstract noun sense that characterizes either suffix overall. In general, the idea that a given word bearing one of these three suffixes simply blocks its rivals does not begin to do justice to the complex interaction both among the suffixes and within individual pairs of words.

words ending in *-ant* and *-ent*, and so is not germane in this particular case. The overall power of *-ness* is revealed in its lack of morphological or other conditioning.

⁵ In case someone is looking for a pattern, the term *eminence* is used as an honorific, but for cardinals of the Catholic Church only, and the expected *eminency* has had little use since the mid 18th century.

It is tempting to see the interactions of rival affixes in terms of synonymy avoidance, of which blocking is a form. Our favorite illustration of synonymy avoidance is an old sociophonetic joke.

Q: What's the difference between a vase [veiz] and a vase [vaz]?

A: Oh, about a hundred bucks.

If blocking and synonymy avoidance were driving the interaction of rival suffixes, then we would expect the rival suffixes to each develop a distinct meaning over time. Remarkably, they do not. There have been attempts to show that the much-discussed rival suffixes *-ity* and *-ness* are no longer synonymous (Riddle 1985), but in environments where *-ity* is productive, for example after *X_v-able*, as in *sustainability* or *likeability*, it has precisely the same range of meaning as *-ness* does elsewhere. Only where it is less productive does it show a difference and then it is precisely what one finds with all less productively formed words—specialization, technical usage, and naming. The standard example is *productivity* as opposed to *productiveness*. We talk of *productivity indices* and *personal productivity practices*. There is even a machine tool company named *Productivity Inc.* In none of these instances would *productiveness* do. Finding circumstances under which only *productiveness* is acceptable is difficult, though the following definition of the term *artificially busy* from *Urban Dictionary* (<<http://www.urbandictionary.com>>) appears to fit the bill: “A state of activity usually reserved for use in the presence of a manager or boss. The activity mimics productiveness without actually serving a purpose.” Here *productiveness* refers only to the state of being productive rather than to some formal measure, which is why *productivity* is at least awkward.

In our own work on rival affixes in English over close to forty years, the only robust example of the members of a set of rival affixes becoming differentiated in meaning is the set *-dom*, *-hood*, and *-ship*. Aronoff and Cho (2001) argue that *-ship* has become specialized to distinguish between stage-level and individual-level attributes. But Lieber (2010a) questions even this case. Based on corpus data she concludes that the three suffixes are frequently interchangeable. This leaves us with no real cases of semantic differentiation in English, the language where this theoretical possibility has been most sought after.

Blocking has proven to be much more successful as a technique in accounting for the interaction of rival realizations in inflection, as opposed to word formation (Brown and Hippisley 2012). Even there, though, problematic nondiscrete rivalries can be found. The English comparative and superlative degree of adjectives, for example, may be expressed either by affixation of *-er* and *-est* or periphrastically with the adverb forms *more* and *most*. Early theoretical accounts of the distribution of the two claimed that the affixal form is found with monosyllables and certain disyllables, with the periphrastic form occurring elsewhere (Aronoff 1976). Less cursory investigation (Graziano-King 1999, Graziano-King and Cairns 2005, Boyd 2007, Gonzalez-Diaz 2008, Mondorf 2009) show that in fact the distribution of the two overlaps, resulting in competition in many individual cases. It may be, then, that even in inflectional systems the distribution of rival forms of realization is not inherently discrete but only becomes so over time.

The study of productivity in word formation over the last quarter century and more has revealed that it is fruitless to conceptualize productivity of word formation in discrete terms, as all or none. Progress in our understanding has been achieved only by assuming that productivity is scalar, entailing the use of statistical methods. Baayen (2003) and Hay and Baayen (2005) provide persuasive extended arguments for this conclusion.

The realization that morphological productivity is a graded phenomenon opens the doors to new and innovative statistical methods. It also allows us to take advantage of the rapidly expanding electronically analyzable data resources that have become available in this period. In the rest of this chapter, we review some of the statistical methods that have been used for quantifying and measuring productivity, along with results. We begin with the best-known electronic corpus-based method, Harald Baayen's measures of productivity based on *hapax legomena*, words that only occur once in a corpus.⁶ We then move to two methods that we have used ourselves. The first takes advantage of digital versions of the *Oxford English Dictionary*, and allows one to trace the productivity of affixes over time. This method was first developed by Anshen and Aronoff (1999). The last method takes advantage of the vast and ever-expanding virtual corpora made possible by the World Wide Web.

5.4 HAPAX LEGOMENA

Linguists have struggled to precisely define what productivity is; quantifying and measuring productivity is, therefore, also problematic. The most useful measures of productivity over the past twenty years have come from the work of Baayen and colleagues, particularly Baayen (1992) and Baayen (1993). These measures, *P* and *P**, center on the notion of the *hapax legomenon*, or a word that occurs only once in a corpus. Baayen's underlying assumption is that there is a strong relationship between hapaxes (as they have come to be called instead of the "proper" Greek plural *hapax legomena*) and productivity.

Baayen's first measure is *P*, which Baayen (1993) calls the Category-Conditioned Degree of Productivity. For a given affix, *P* is defined as:

$$P = n_i / N$$

where n_i represents the total number of hapaxes containing the affix, and *N* represents the total number of tokens containing the affix. This measures the "growth rate" (Baayen 1992) of the affix: the probability that an encounter with a word containing the affix reveals a new type.

⁶ The term *hapax legomenon* 'read once', sometimes plain *hapax*, is Greek and originates in the scholarly study of the Bible, where the meaning of a word that only occurred once in the received text might be especially difficult to discern, making such words of special interest.

Ideally, n_i would precisely represent all individual word types in a corpus that were productively derived, regardless of the number of times the word occurs in the corpus; however, it is certainly not feasible, and probably impossible, to systematically test whether a given token in a corpus was created productively or came from that speaker's lexicon. Both P and P^* (which we will discuss shortly) must rely on the assumption that hapaxes are a good representative of productive word formation; indeed, Baayen (1993: 189) explains that the probability of encountering neologisms "is measured indirectly" via the counting of hapaxes, and that not all hapaxes are neologisms, and vice versa. Given this, it is crucial that it be true that, if a token occurs only once in a corpus, it is proportionately more likely to be productively formed, and, conversely, if a word is productively formed, it is proportionately more likely to occur only once in a corpus. Intuitively, this seems like a sensible assumption, but it is difficult to prove; to do so would require, at the very least, a precise, agreed-upon set of criteria for categorically judging an occurrence of a word to be productively formed or not productively formed. Paradoxically, the need for measurements like P and P^* arise precisely because this cannot be accomplished. Instead, Baayen supports the use of P and P^* as measures of productivity by analyzing their predictions: do the measurements produced by these methods yield results that correlate with our intuitions about the affixes in question?

Baayen (1992) assesses the validity of P by comparing rival suffixes such as English *-ity* and *-ness* using the CELEX database. Of the two, *-ness* is qualitatively regarded as much more productive than *-ity*, although there are a large number of established *-ity* types.

As shown in Table 5.1, Baayen's P measure produces a value of 0.0007 for *-ity* and 0.0044 for *-ness*, even as the number of types (405 and 497, respectively) is very close between the two. Further, both P values are higher than the P value of 0.0001 for simplex nouns in the corpus (which, by definition, are not formed through productive processes).

To evaluate global productivity, Baayen suggests considering both P and V together, where V is the number of individual word types in the corpus (the vocabulary size). Differences in V reflect the extent to which relevant base words have been used, while differences in P relate to differences in extent that remaining base words can be used to create neologisms.

Table 5.1 Comparing the productivity of *-ity* and *-ness*

Affix	Tokens	Types	Hapaxes	P
simplex nouns	2,142,828	5,543	128	0.0001
<i>-ity</i>	42,252	405	29	0.0007
<i>-ness</i>	17,481	497	77	0.0044

Source: Adapted from table 2 in Baayen (1992).

Table 5.2 P^* value comparison

Category	$P^* \cdot h_1$ (a.k.a. hapaxes)	Qualitative judgment of productivity
simplex nouns	256	–
-ness	77	+
-ation	47	+
-er	40	+
-ity	29	+
-ment	9	±
-ian	4	±
-ism	4	+
-al	3	±
-ee	2	±

Source: Adapted from table 3 in Baayen (1993).

Baayen's second measure, the Hapax-conditioned Degree of Productivity, P^* , is defined as the following:

$$P^* = n_i/h_i$$

Again, n_i is the total number of hapaxes with a given affix, while h_i is the total number of hapaxes across all types in the corpus. This measure predicts the likelihood that any new word that one encounters will contain the affix, which, according to Baayen (1993: 193) "can also be viewed as measuring the relative contribution of a given morphological category to the overall vocabulary growth."

P^* is tested in Baayen (1993) in similar fashion to P , by judging its predictions. Because h_i is the count of total hapaxes in the corpus, the denominator in the expression n_i/h_i is the same for all suffixes; therefore, when comparing P^* values in a given corpus, we are, in effect, simply comparing the number of hapaxes occurring for each suffix.

Returning to *-ity* and *-ness*, we see in Table 5.2 that there are 29 *-ity* hapaxes and 77 *-ness* hapaxes, which is again in line with our intuition that *-ness* is more productive. With P^* , however, it is not possible to compare values to a baseline of simplex nouns, as the P^* value for this category is very high.

Both P and P^* measurements are dependent on the size (N) of the corpus. The number of hapaxes in a corpus is a decreasing function of N ; ultimately, the rate of increase in the number of hapaxes slows as the size of the corpus increases. This means that comparing measurements across corpora is problematic.

Hay and Baayen (2002) show a link between parsing and productivity—namely, Baayen’s *P* measurement of productivity. For words containing a given affix, Hay and Baayen plot the frequency of the derived forms against the frequency of bases of those forms. In forms where $x = y$, the frequency of the base is equal to the frequency of the derived form; Hay and Baayen call the line $x = y$ the *parsing line*. Those forms that are plotted below the parsing line are words that are more frequent than their bases (e.g. *illegible* is more frequent than *legible*). Forms above the parsing line have bases that are more frequent than the derived forms. Hay and Baayen claim that those words falling below the parsing line are more likely to be accessed as whole words, rather than component parts, while words above the line are more likely to be decomposed and the affixes, therefore, used productively. They then calculate the *parsing ratio* for a given affix, that is, the proportion of words that appear above the parsing line. Hay and Baayen find that *P* is a strong predictor of this parsing ratio, providing further support for the validity of *P* as a measurement of productivity.

5.5 DICTIONARIES

Detailed historical dictionaries, such as the *Oxford English Dictionary* (*OED*), can provide a rough, but nonetheless insightful, diachronic survey of productivity. They allow us to address a different question from most studies of morphological productivity. Rather than what it means for a given morphological structure to be more or less productive than another, they allow us to study how a given morphological structure has become more or less productive over time.

Any dictionary is subject to the biases of the editors. Dictionaries of standard written languages tend to favor works of well-regarded authors as major sources of citations. Thus, one can neither assume that all “existing” words are in the dictionary, nor that all words in the dictionary are currently “existing” words. However, it is probably impossible to compile such a list.

Only in the kind of ideal world that contains ideal speaker-listeners can we hope to find a list of existing words. It follows that the methodologically practical assumption of the equivalence of the word-list of any reference work or set of reference works and the set of existing words is inevitably flawed. (Bauer 2001: 36)

Anshen and Aronoff (1999) use the *OED* on CD-ROM to investigate the birth and death of borrowed suffixes in English. Using the software’s advanced searching tools, one can search for words matching certain criteria, such as all words ending in the suffix *-ity*. Each word’s entry contains (among other things) definitions, etymological information, and citations. The date of first citation can be used as an approximate indicator of when a word came into use, while the etymology makes it possible to determine the likely

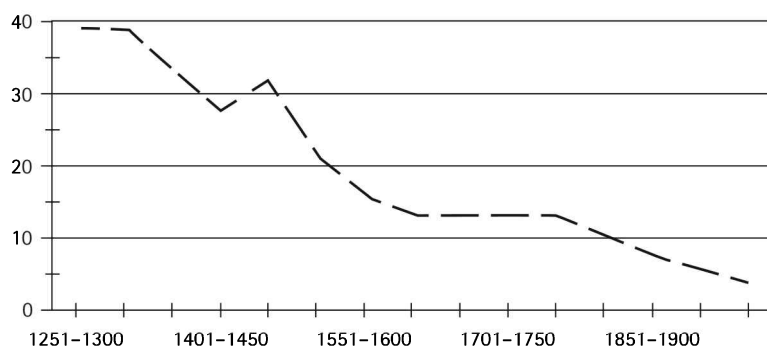


FIGURE 5.1 French borrowings as a percentage of all new words

Source: from Anshen and Aronoff (1999).

language of origin. The latter piece of information can be used to categorize a word as being borrowed or native to English.⁷

Grouping these dates of first citation into bins by century or half-century, one can graph the number of words cited for the first time in each time period, giving an approximation of how productive an affix has been over time. In Figure 5.1 and Figure 5.2, Anshen and Aronoff show the birth of *-ity* as a productive suffixation pattern in English. Figure 5.1 illustrates the gradual decline in the borrowing of French words into English. Over the same time period, we see in Figure 5.2 that an increasing percentage of new *-ity* words were being derived in English. Indeed, during the 19th century, 937 new *-ity* words were derived, while only 35 were borrowed.

We see a gradual increase in *-ity* derivations, even as borrowings decrease; this is the birth of *-ity* as a productive suffixation pattern in English. Once a sufficient number of borrowings had entered into the language, speakers could then abstract out the suffix and generalize a pattern. Anshen and Aronoff then use this same method to compare the differing fates of productive *-ment* and *-ity*. While *-ity* is a productive suffix today, *-ment* has all but fallen out of use, except as a fossilized component of established words.

In Table 5.3, Anshen and Aronoff track the number of new derived *-ity* and *-ment* forms entering English. Here we see a strong decline in new *-ment* words beginning in the 17th century, while *-ity* generally holds strong into the present day. This decline in *-ment* derivations coincides with a change in the number of new verbs entering into English, as shown in Figure 5.3. Since *-ment* is dependent on new verbs for productivity, a decline in potential hosts should have an impact on its performance as a productive

⁷ Defining the language of origin for a word is by no means trivial, and the origins chosen by the *OED* are subject to interpretation. For example, if an affixed word's first citation in English comes later than the first citation of the same word in French, this does not guarantee that the English speaker did not simply derive the same word natively. However, it may be impossible to determine such a thing definitively.

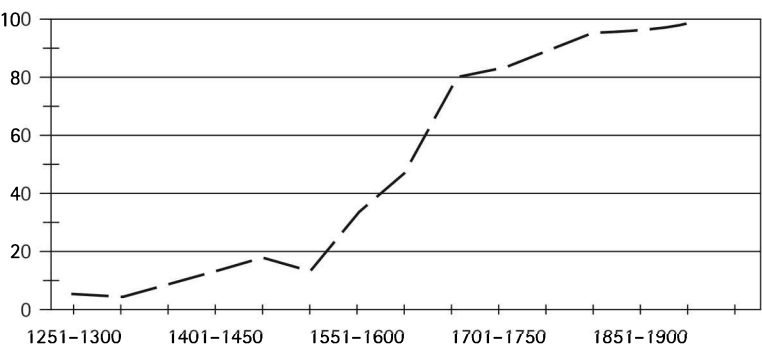


FIGURE 5.2 Derived *-ity* as a percentage of all *-ity* words

Source: from Anshen and Aronoff (1999).

Table 5.3 Derived forms for *-ment* and *-ity*

Half-centuries	Derived <i>-ment</i>	Derived <i>-ity</i>
1251-1300	6	1
1301-50	10	1
1351-1400	19	11
1401-50	15	11
1451-1500	37	16
1501-50	60	22
1551-1600	174	64
1601-50	217	206
1651-1700	76	241
1701-50	40	108
1751-1800	37	177
1801-50	158	435
1851-1900	142	502
1901-50	26	298
1951-2000	4	179

Source: from Anshen and Aronoff (1999).

pattern; on the other hand, *-ity* relies on adjectives, which enter English in greater numbers during the same time period, and it thrives productively.

Lindsay and Aronoff (2013) improve on this claim by normalizing the *OED* data to account for the variable amount of source material from century to century. After

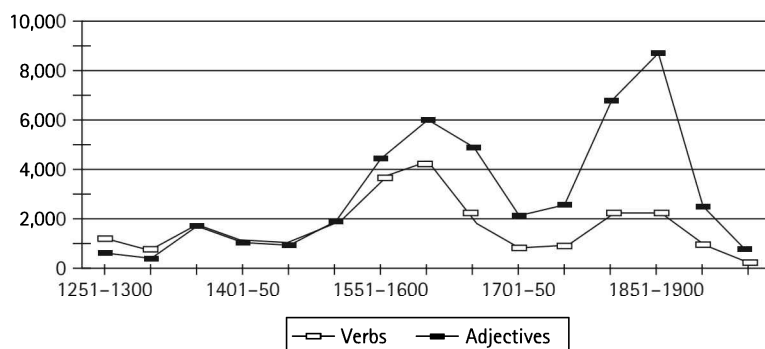


FIGURE 5.3 Number of new English verbs and adjectives

Source: from Aronoff and Anshen (1999).

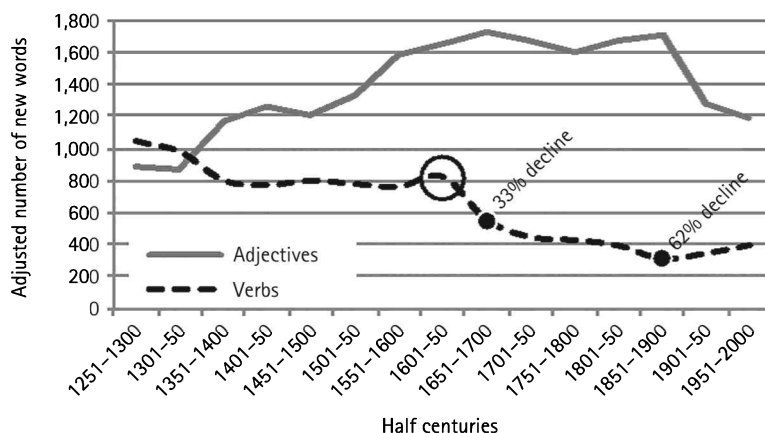


FIGURE 5.4 New adjectives and verbs entering English, showing a rapid decline in the relative number of new verbs beginning in the 1600s

Figure 5.3 is adjusted,⁸ the difference between new verbs and new adjectives becomes much clearer (Figure 5.4). Likewise, the divergence between *-ity*'s productivity and *-ment*'s productivity is also much more pronounced (Figure 5.5, compared to Table 5.3). Thus, we see that historical dictionaries like the *OED* provide a practical means of analyzing the relative productivity of suffixation patterns diachronically. This information can be used to track the emergence (and death) of productive processes and identify factors that may have influenced their fates.

⁸ The value for the number of words in a given half-century is proportional to the total words in the *OED* for that time period:

$$\text{adjusted number of words} = (\text{number of words} / \text{total words}) \times 10^5$$

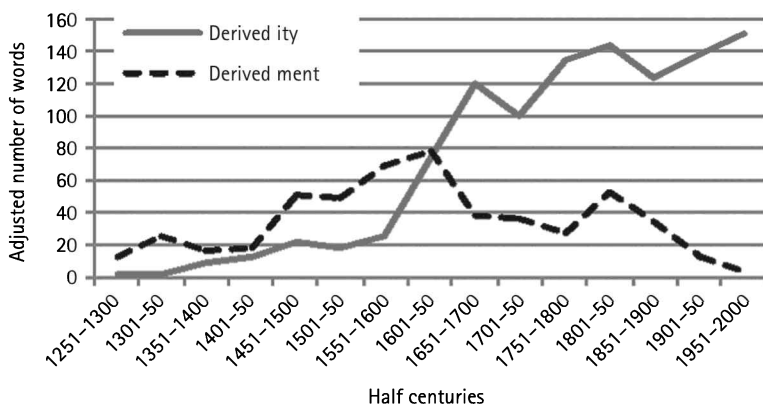


FIGURE 5.5 New derivations of *-ity* versus *-ment* over the past 750 years

5.6 THE WORLD WIDE WEB AND PRODUCTIVE RIVALRIES

Rivals *-ic* and *-ical* are both productive today in spite of their mutual dependence on the same pool of stems and no distinguishing semantic differences.⁹ Why do some rival patterns seem to stabilize and coexist, while others do not? Lindsay and Aronoff (2013) view languages as self-organizing in a manner similar to biological systems; languages are complex, continuous systems that change through numerous smaller interactions, a phenomenon known as glossogenetic evolution (Hurford 1990, also discussed in Steels 1997 and Fitch 2010, among others).

Lindsay and Aronoff (2013) also use another resource, the world wide web, to examine synchronic productivity of suffix rivalries in English. To accomplish this, they use statistical estimates from search engine results—in this case, the Google Search engine, using the Google Search API.¹⁰

One must be cautious when incorporating Google Search's Estimated Total Matches (ETM) into a measurement of usage. While Google is a vast and freely-available resource, it is also “noisy”; that is, individual results contain false positives due to typos, non-native speech, spam, the lack of part-of-speech tagging, and so on. Furthermore, ETM results represent the number of pages a string

⁹ While word pairs like *electric* and *electrical* have different meanings, these differences are not generalizable; the difference between these words bears no resemblance to the difference between e.g. *historic* and *historical*.

¹⁰ As of January 2012, the Google Search API has been discontinued for all purposes, including academic research. Querying Google for results is still technically possible, but much less practical, due to the constraints made by Google on query frequency.

is estimated to appear in, not the number of occurrences. (Other discussion of such considerations can be found in Hathout and Tanguy 2002, among others.) For these reasons, it is important that little weight is placed upon the actual raw numbers themselves (only relative differences should be considered) or upon any individual word pairs. A broad investigation of suffixes mitigates many of these concerns when dealing with single words, regular inflection patterns, and a large number of stems (Lindsay and Aronoff 2013).

To gather the data, first, a list of suitable words must be generated in order to feed them into Google Search.¹¹ Using basic regular expression matching (along with some manual filtering), we can identify all words ending in either *-ic* or *-ical* in Webster's 2nd International dictionary.¹² The suffixes of these words are then stripped off, leaving bare stems; duplicate stems are discarded. In the case of *-ic/-ical*, this yielded 11,966 unique stems that take *-ic*, *-ical*, or both suffixes.

Next, each stem-suffix combination is automatically queried in Google Search as a literal string (e.g. *biolog + ic*, *biolog + ical*) and the ETM value is returned and recorded in a database. ETM values are then compared and analyzed. In Tables 5.4 and 5.5, we see a sample of ETM values for various *-ic/-ical* pairs. In some cases, both *-ic* and *-ical* have a substantial number of tokens (Table 5.4), though in the majority of cases, one suffix yielded far more results than the other (Table 5.5). Overall, 88.5% of pairs differed by at least one order of magnitude. By comparing ETM values for each form for a given stem (e.g. *biolog-ic* and *biolog-ical*), the assumption is that the more productive suffix will tend, over a large number of comparisons, to have a higher ETM value more often than the less productive suffix.

Between *-ic* and *-ical*, *-ic* was found to be the "winner" in 10,613 out of 11,966 pairs. However, *-ic* was not preferred in all domains. Lindsay and Aronoff systematically examined all neighborhoods appearing on the right-edge of the list of stems. For example, one could look at the final letter of each stem (neighborhood length 1) and find that there are 4,166 stems ending in *t*, or look at the final two letters (neighborhood length 2) to see that there are 1,129 stems ending in *st*. If one continues this process for all possible combinations, generalizations begin to emerge.

Naturally, there is clustering in certain neighborhoods, with the largest groups usually (but not always) coinciding with traditional morpheme boundaries (e.g. *graph*). Most of these groups also favor *-ic* over *-ical*; however, there is one significant exception: stems ending in *olog* (of which there are 475) favor *-ical* over *-ic* by a ratio of 6.42 to 1 (Table 5.6).

¹¹ At one time, it was possible to use regular expression matching in search engine queries; for example, Hathout and Tanguy (2002) created the WebAffix tool for these types of advanced queries using the AltaVista search engine. Unfortunately, advanced regex matching, if present at all, is severely restricted in present-day search engines. Thus, querying must be done without the direct use of regex matching.

¹² This was chosen as a source, in part, because it was freely available in digital form.

Table 5.4 Sample Google ETM counts for high-frequency doublets

Stem	-ic count	-ical count	ratio (-ic/-ical)
electr-	325,000,000	218,000,000	1.49
histor-	133,000,000	258,000,000	0.52
numetr-	23,900,000	37,200,000	0.64
logist-	13,000,000	5,850,000	2.22
asymmetr-	10,400,000	6,410,000	1.62
geolog-	7,980,000	22,800,000	0.35

Source: from Lindsay and Aronoff (2013).

Table 5.5 Sample Google ETM counts for high-frequency singletons

Stem	-ic count	-ical count	ratio (-ic/-ical)
civ-	90,000,000	2,220	40,540
olymp-	73,300,000	1,130	64,867
polyphon-	32,800,000	869	37,744
sulfur-	10,600,000	0	—
mathemat-	1,740,000	48,900,000	3.56×10^{-2}
typ-	421,000	158,000,000	2.66×10^{-3}
theolog-	71,300	18,100,000	3.94×10^{-3}
post-surg-	287	1,090,000	2.63×10^{-4}

Source: from Lindsay and Aronoff (2013).

Table 5.6 -ical is productive in stems ending in olog

	Total stems	Ratio	-olog stems	Ratio
<i>Favoring -ic</i>	10,613	7.84	74	1
<i>Favoring -ical</i>	1,353	1	401	6.42
<i>Total</i>	11,966		475	

Source: from Lindsay and Aronoff (2013).

Here we see a possible explanation for the ability of rivals *-ic* and *-ical* to coexist productively in a competitive ecosystem. While *-ic* is strongly preferred overall, *-ical* is not being driven out of the system because there exists a coherent subdomain of sufficient

size in which *-ical* is preferred. Without this niche, we should not expect *-ical* to survive as a productive entity (see Lindsay and Aronoff 2013 for further discussion).

While Google Search estimates cannot be employed for all linguistic (or even all morphological) investigations, in certain applications these data can provide valuable insight. Although the research program for Google Search has been terminated, the Google Books corpus remains freely available. Version 2 of the English language corpus (released in 2012) provides information on the date of citation, as well as part of speech tagging, on millions of books and publications.

CHAPTER 6

METHODOLOGICAL ISSUES IN STUDYING DERIVATION

ROCHELLE LIEBER

6.1 INTRODUCTION

It seems safe to say that since the days of the American Structuralists discussions of methodology have not been at the forefront of the study of word formation. In the scientific climate of Structuralism, it was a major goal to be able to build a theory by applying a set of analytic operations to observable data; this sort of attention to methodology would ostensibly make linguistics scientific. The empiricist climate of the times lent itself to the search for so-called “discovery procedures” whereby the grammar of a language could be constructed from a corpus of data. The procedure was largely bottom up, starting with the delineation of phones, their assignment to phonemes, and then the assignment of phonemes to morphemes, and so on (Harris 1955, Newmeyer 1986).

With the advent of the generative movement, discussions of methodology became at best tangential to what was perceived as the real work of linguists: constructing grammars that model the mental representations native speakers have of their language. The attempt to find “discovery procedures” was widely acknowledged to be misguided. When morphology resurfaced in the generative tradition as a legitimate subject of study in its own right, it was tacitly assumed that whatever methods syntacticians and phonologists used in obtaining data were equally suitable for obtaining morphological data. Those methods—especially for syntacticians—largely involved data generated on the basis of the intuitions of native speakers, often the linguist herself.

In this chapter I will suggest that the issue of methodology is perhaps more pressing in the study of morphology than in other areas of linguistics, and most critical in the study of derivational morphology. I will discuss the strengths and weaknesses of various ways of gathering data, including the use of intuitions, dictionary data, texts and corpora, and psycholinguistic experimentation. My argument will be that reliance on self-generated data and personal intuitions alone is particularly problematic in studying derivational

word formation, that the theorist must be open to a wide variety of approaches, and that cross-fertilization among theoretical, corpus-based, and psycholinguistic experimental approaches is very much to be desired. This is not, then, a call for a return to the era of discovery procedures, but a suggestion that some degree of eclecticism in looking at derivational morphology will lead to a far richer understanding of the mental lexicon.

6.2 THE ROLE OF SELF-GENERATED DATA AND NATIVE SPEAKER INTUITION

The American structuralist Zellig Harris (1955) was clear that linguistic analysis should be done on the basis of a corpus of data. Nevertheless, given the lack at the time of suitable corpora and of computers capable of manipulating corpus data, this was not in fact the way that structuralists studied language. Consider, for example, Harris's description of his own procedure in segmenting utterances into morphemes:

The procedure requires a large number of associated utterances sectionally identical with U [some particular utterance—RL], some in their first phoneme, others in their first two phonemes, and so on. We could draw these utterances from some written corpus; but the corpus would have to be prohibitively large if we are to be able to find in it, for any U we choose, enough associated utterances for each *n* of U. The only practicable way of finding the required utterances is to elicit them from an informant, i.e. to ask him for any utterances beginning with /h/, then for any utterances beginning with /hi/, and so on. (Harris 1955: 194)

As Harris notes, it might be desirable to base one's analysis on a corpus of real-life utterances, but this was not actually a practical possibility at the time. To the extent that a corpus could be used, the notion of corpus was not our contemporary one, but a body of language elicited from a speaker.

Post-structuralist morphologists (myself among them) have typically not given much thought to the issue of where we get our data; following the lead of syntacticians, we have at least until recently tended to base our analyses on examples drawn from our own knowledge of a language, sometimes supplemented with data from dictionaries, grammars, reference works, and the occasional example we come across in our own experience (reading, radio, TV and other media, etc.). Self-generated data are certainly good enough to suggest the shape of a model of derivation that is primarily concerned with the basic internal structure of words, but they nevertheless raise a number of issues.

The first issue is that examples that come to us off the tops of our heads tend to be words that are item-familiar and high frequency. Such forms are likelier to have lexicalized meanings and perhaps bias the morphologist towards believing that derivational patterns are more idiosyncratic than they might in fact be. Neologisms, or at least item-unfamiliar and low-frequency forms often tell us more about derivational patterns,

since they typically are both formally and semantically more compositional. But low frequency and item unfamiliar forms are not the ones that are likely to come to mind.

We can of course try to generate neologisms ourselves, but this practice leads to a second issue, namely the status of negative intuitions. It is not uncommon in the literature to see comments to the effect that a particular word or a whole pattern of derivation in a language is not possible. Consider, for example, the following claim from my own dissertation (Lieber 1980: 115): “*Re* and *un* could only attach to verbs involving a change of state, and *kill* is not such a verb. But *rekill* and *unkill* sound far less deviant than words like **unpeace* and **refusity*.” Interestingly, it is possible to find attestations for both *rekill* and *unkill*, as well as for *unpeace* (the former two from Google Books, the latter from the Corpus of Contemporary American English (COCA)) none of which sounds particularly odd in context:

- (1) *rekill*: “If you don’t care for it, you’d best hurry off. Lazarus aims to **rekill** you.”
(Richard Laymon, *Savage* 2007: 290)
- unkill*: “They kill, then **unkill** me, and I emerge changed.” (*The Best Buddhist Writing* 2009: 209)
- unpeace*: “A wandering eye, a divided heart, a stifled cry, that is my repeated experience of **unpeace**, of restlessness.” (*Cross Currents* 1990)

It is true that these words are infrequent occurrences in the corpora—we have to go to a corpus of 155 billion words to find *rekill* and *unkill*—but there they are.

Selkirk (1982: 34), in her treatment of synthetic compounds provides us with a second example of the problem with negative intuitions. There she claims that, “The SUBJ argument of a lexical item may not be satisfied in compound structure,” thus ruling out compounds like *girl swimming* or *kid eating* (her examples, the latter referring to children’s habits rather than cannibalism). While intuition might again suggest that those particular words are a bit odd, other compounds in which the first element is interpreted as subject can easily be found and seem intuitively quite ordinary when seen in context; for example, the compound *airline hiring* (meaning ‘hiring by airlines’ rather than ‘hiring airlines’) is attested in COCA, and seems quite unremarkable.

Another example can be found in Di Sciullo and Williams (1987: 39). Di Sciullo and Williams claim that nouns derived by conversion from verbs cannot take arguments (**the hit of Bill*, **the kick of Bill*). Again, although there may be something to their intuitions, comparable constructions are not hard to find in corpus data and again are relatively unremarkable in context: the phrases *Edmund Hillary’s climb of Mount Everest* and *his stop of this car* can be found in COCA.

One particular theoretical construct, namely “blocking” has largely been based on the sort of intuition that if item *x* occurs, item *y* cannot also occur. For example, with respect to nominalization, intuitions might tell us that verbs tend to have specific nominalizations and that other conceivable ones are blocked. For example, the following nominalizations are all familiar: *display_N*, *disregard_N*, *disruption*, *fluctuation*, *revision*, *cessation*, *omission*, *discharge*, *conjunction*, *revelation*. We might be inclined then to say that forms

like *displayal*, *disregardance*, *disrupture*, *fluctuance*, *revisal*, *ceasement*, *omitment*, *dischargement*, *conjoinment*, and *revelment* are blocked, at least if they are meant to be synonymous. Yet all of these nominalizations are attested in COCA, and at least some of them seem to be used with exactly the same meaning as the familiar nominalization (see Bauer et al. 2013). Granted they are rare, but what are we to say about rare occurrences? From one point of view it is precisely these item-unfamiliar forms that give us evidence of productivity in derivation. Indeed, one well-accepted quantitative measure of productivity, Baayen's (1989) measure P relies on the proportion of items with a particular corpus that occur with a frequency of one, so it would seem that we might dismiss such examples at our peril.

Self-generated data also tend not to give us the full picture when we are studying semantic patterns in derivation. The case of denominal verbs in *-ize* is instructive. Broadly speaking, *-ize* might be thought of as a causative prefix. Where *-ize* attaches to an adjective A, we typically get the meaning 'cause to become A' or 'make A'; so *legalize* means 'to make legal.' But scrutiny of a larger set of data shows that there is much more to be said. As detailed in Plag (1999) and Lieber (2004: 77), there is a wide range of meanings that are attested in words derived with *-ize*, among them the following:

(2) Semantics of *-ize*

'make x', 'cause to become x'	standardize, unionize
'make x to to/in/on something'	apologize, texturize
'make something go to/in/on x'	hospitalize, containerize
'do/act/make in the manner of x'	Boswellize, despotize
'do x'	philosophize, theorize
'become x'	oxidize, aerosolize

Not all denominal verbs in *-ize* are "causative" in the same way, since the base noun can be interpreted either as theme or goal, and further, not all verbs in *-ize* are causative at all: as the examples above suggest, some are activity verbs and a few are purely inchoative. Intuitions alone cannot tell us how strong each of these patterns is and it seems safe to say that we run the risk of underestimating the extent of affixal polysemy if we rely on self-generated data.

6.3 DICTIONARIES, GRAMMARS, AND THE PRESCRIPTIVE TRADITION

Morphologists have, of course, frequently made use of dictionaries, backwards word lists, and reference works to supplement their intuitions. For morphologists working on English we have the *Oxford English Dictionary* (OED), now searchable in quite sophisticated ways online, as well as Lehnert (1971), a useful backwards word list, and such reference works as Marchand (1969) and Jespersen (1942). While such

supplementary data are valuable, and are often more than adequate to give us a good idea of the extent of affixal polysemy, they are inevitably biased towards idiosyncratic forms. By their very nature, dictionaries look for a certain threshold of usage before a word comes to be recorded, so the low frequency and item unfamiliar forms that stand to tell us the most about derivational patterns are often not to be found. Moreover, the more compositional the meaning of a derived word, the less it actually needs to be recorded in the dictionary.

A related issue is that people—even morphologists—have a tendency to take dictionaries and reference works as prescriptive, even if it is not the intention of those works to be so. The fact of being recorded in a dictionary gives a word a degree of legitimacy that it might not otherwise have. Even morphologists may be prone to the naïve view that if a form is not in the dictionary (for example, neither *displayal* nor *omitment* are to be found in the *OED*), it is not a “real” word. This is probably likeliest with novel forms derived with less productive affixes. For example, native speakers of English would probably not have trouble agreeing that any new adjective adding *-ness* would result in a legitimate word, even if it does not occur in the dictionary; the *OED* lists *colostral* in its new entries for September, 2011, and *colostralness* would likely be acceptable to native speakers, although it is not recorded in the *OED*. However, with affixes whose productivity is in dispute (for example, *-al* or *-ment*) we might be more inclined to take absence from the dictionary as significant. The most obvious conclusion we should draw is that the presence of a word in a dictionary or other reference work can be useful to the morphologist, but that absence cannot be taken to mean impossibility.

6.4 LITTLE STUDIED LANGUAGES

Interestingly, some of the issues I have raised above, although not clearly on the radar of morphologists working in the generative tradition on well-documented languages, have been more prominent in the literature on language documentation. For example, Haviland (2006: 129) begins his consideration of lexical documentation with the issue of when “enough is enough,” that is, at what point in doing fieldwork we can assume that we have discovered enough of what there is to discover about the lexicon of a language (his answer: “seemingly never”); although Haviland is mostly concerned with basic documentation of the lexicon rather than with morphology per se, we can extend his concern to documentation of derivation, as this is often the primary way languages add to their lexical stock. How do we make sure we’ve gathered all the facts? How much data do we need?

The discussion frequently hinges, not surprisingly, on the relative weight to be given to elicited versus “naturally occurring” data. The advice in Payne (1997) is especially perceptive. The main purpose of his book is to expose students to the types of morphology (both derivational and inflectional) they might encounter in unfamiliar languages, but in a very brief chapter at the end of the book Payne discusses the relative merits of

elicitation of examples as opposed to finding examples in text. His thoughts are worth quoting at length:

Good text data are uncontrolled, open-ended, and dynamic. A text will contain forms that never appear in elicitation. It will also contain forms that appear in elicitation, but in sometimes obviously and sometimes subtly different usages. There is much idiosyncrasy in text. That is, forms are used in novel ways in order to accomplish very specific communicative tasks. Sometimes these are referred to as “nonce” usages. For example, a sentence like *He psycho-babbled away our two-hour appointment* might arise in a particular communication situation, even though the verb *to psycho-babble* is probably not a part of the lexicalized vocabulary of most English speakers. One wonders how such a sentence could possibly be elicited! Such idiosyncrasy in text is more common than one might expect and often provides great insights into speakers’ ways of thinking and conceptualizing their experience. (Payne 1997: 367)

We might add that data of this sort are invaluable in gauging productivity and semantic nuance in derivation.

Payne recommends using elicitation to arrive at an “inventory of derivational morphology (which derivational operations apply to which roots, etc.)” (1997: 368), but in time progressing to use texts almost exclusively for studying “lexical semantics (determining the nuances associated with various lexical choices, including derivational morphology. . .” (1997: 369). The insight here is that relying exclusively on elicitation gives a useful but limited picture, and is especially limited with regard to exposing semantic nuance. The upshot is that textual data can add immensely to our understanding of derivation.

6.5 THE USE OF CORPORA

The advantages of using textual data for studying better-studied languages have only recently become apparent. In syntax, the work of Bresnan et al. (2007) and Bresnan and Ford (2010) on the dative alternation is instructive (that is, *give NP₁ to NP₂* vs. *give NP₂ NP₁*). Bresnan and Ford (2010: 170) point out that intuitions with respect to syntactic patterns are sometimes inconsistent with corpus data, noting that, “. . . reported cases of non-alternation based on intuitive judgments of decontextualized examples are surprisingly inconsistent with actual usage. . .” As I have indicated above, the same seems to be the case with respect to derivational morphology; nevertheless, it is only recently that morphologists have begun to mine large corpora systematically in studying derivation. This should perhaps not be surprising. Finding derived words in text can be a scattershot affair, and the probability of finding forms that give new insight into well-studied languages might seem relatively low. But the larger the corpus, the better our chance of finding those “rare events” that I would argue turn out to be the most useful to the morphologist studying derivation.

It is only relatively recently that corpora have become available that are both large enough and easily enough manipulated to overcome the problems with basing theories on self-generated or dictionary data. For English, the British National Corpus (BNC), with 100 million words of spoken language and written text from 1980 to 1994, has been available for some time, but only recently with the user-friendly interface provided by Mark Davies' website (<<http://corpus.byu.edu/>>). Even more useful is COCA with 450 million words (at the moment of this writing—COCA is added to yearly), with both spoken and written language spanning a wide range of genres. Google Books, with 155 billion words, is a vast source of written language, but less conveniently searchable.

Mining data from these sources can still be a laborious and time-consuming enterprise: for example, to find all attested instances of words with a particular suffix, say *-ity*, requires doing a wildcard search (that is, searching for **ity*) and painstakingly cleaning the resulting list by deleting “junk”—misspelled items, items with weird punctuation, and items that end in the string *ity*, but not in the suffix (*city*, *pity*, *uppity*, etc.). As a next step, we must further make sometimes difficult decisions whether to count a word as one having the affix in question or not (for example, among the *-ity* forms do we consider forms on bound bases like *acuity* and *levity*? what do we do about jocular forms like *craposity*?). Although the process of obtaining the data is tedious, the payoff can turn out to be substantial, however.

First, by observing “rare occurrences” we can begin to see that processes that do not intuitively seem terribly productive in fact do give rise to apparently novel forms. In preparing *The Oxford Reference Guide to English Morphology* (Bauer et al. 2013), my co-authors and I were constantly surprised by areas of English derivation that were continuing to produce new words. For example, nominalizers like *-ment* and adjective-forming suffixes like *-ory* and *-ous* show modest numbers of novel forms, and even with affixes for which there are relatively few types overall (e.g. *-ive*) forms can be found in the corpora that are not recorded in the *OED*. We can, of course, question what weight we should give to rare novel forms, whether they should be generated via rules or by some sort of analogy (whatever we might mean by analogy), but it is hard to argue that they should be ignored entirely.

Further, because the items gathered from corpora such as BNC and COCA can be viewed in context, they allow us to explore the semantic subtleties of affixes in ways that were not formerly available systematically. Again, an example might be instructive. Whereas it is generally the case that the derivational suffix *-er* is an agentive and instrumental affix, or more broadly a subject-oriented personal affix (Booij 1986, Rappaport-Hovav and Levin 1992), we can learn from corpus data that its semantics is more complex. Consider, for example, this citation from COCA:

- (3) *Outdoor Life* 2005: I had taken bears before and had been hunting for several years for a truly outstanding bear, and here one was standing broadside at 20 yards. I didn't have to think twice about this bear. It was a **shooter**.

In this context, the *-er* form is not an agent or instrument noun, nor is it even subject-oriented. Rather, it is a patient noun. Why use *-er* here rather than the usual

patient noun suffix *-ee*? First, it has been pointed out that *-ee* nouns typically denote sentient beings (Barker 1998). We have a tendency to coin *-er* patient nouns when the referent is conceptualized as non-sentient. But further, as this example suggests, what is meant here is not just that something undergoes the act of shooting, but further that it is *meant* to do so. This nuance is only rarely to be found in lexicalized forms (*loaner*, *keeper* are two that come to mind), but probably not all that unusual in nonce forms. Close analysis of corpus data can show us the range of senses for any given affix, and returning to the issue of productivity, can also suggest the extent to which any given reading is productive.

6.6 PSYCHOLINGUISTIC METHODOLOGY (REACTION TIME, APHASIA, NEUROIMAGING)

As someone whose main contribution to morphology has been in the area of theory, I am not well-equipped to give a detailed critique of the methodology that has been used in psycholinguistic study of derivation. Chapter 7 will look at this tradition in some detail. Here I will give only a broad-brush outline in order to highlight the point that the questions that have driven psycholinguistic work on derivation have been somewhat different than those that have motivated theorists.

At least since the mid-1970s there has been a robust tradition of psycholinguistic research delving into the processing and production of complex words. Theorists have generally assumed that there are rules of some sort that give rise to derived words, and have concentrated on questions of representation: for example, whether rules should be morpheme-based or realizational, whether derivation should be subject to the same sort of rules as inflection; how we should measure and model productivity; whether blocking exists and if so what it is. Psycholinguists, on the other hand, have concentrated on questions concerning lexical access and word recognition: whether derived words are processed as wholes—that is retrieved from storage—or decomposed on line on the basis of rules that are represented in the mental lexicon, or both. This has been referred to as the “words and rules” or “computation versus storage” question. Questions of representation (if there are rules what sort of rules? if not, how do we account for generation of new words?) have not been at the forefront.

One basic line of research has involved lexical decision experiments, that is, experiments that measure the reaction times of subjects to simple, derived, compound, or inflected words (Taft and Forster 1975, Butterworth 1983, Schreuder and Baayen 1994, Baayen et al. 1997, McQueen and Cutler 1998, Bertram et al. 2000b, Rastle et al. 2004, de Vaan et al. 2007, among many others). With regard to derived words in particular, psycholinguistic researchers have manipulated not only the degree to which the meaning of individual derived words is compositional, but also the frequency of the base, the

frequency of the derived word itself, the relative frequency of the derived word vis-à-vis the base, and the family size of individual exemplars (the summed frequency of all words derived with a particular base). Experiments often test the effects of “priming” on word recognition, that is, the extent to which exposure to a related word influences the time it takes for a subject to decide if a later-presented word is in fact a word or not. Priming itself can be manipulated in several ways, with presentation of the prime being overt or covert (so-called “masked priming,” where the prime is presented to subjects so briefly that subjects are not conscious of being exposed to it), and if overt what the time lag is between prime and target. For an overview of this methodology, see Hay (2001) and Chapter 7 of this volume.

Experimental studies have also studied the speech of aphasics and their performance in repetition tasks (Badecker and Caramazza 1989) or compared aphasic and normal subjects in lexical decision tasks (Hagiwara et al. 1999). Studies have increasingly also made use of sophisticated imaging techniques to probe the processing of complex words in real time; see for example Solomyak and Marantz (2009, 2010), Lewis et al. (2011). As in the lexical decision literature, the aim in these studies is largely to probe the balance of storage vs. computation, with the added goals of tracking the time course of lexical access and the localization of computation in the brain.

These lines of research have yielded a variety of interesting, but frequently conflicting results (see Hay 2001 for extensive discussion), with the balance of evidence seeming to point towards a combination of storage and computation. There has been little consideration, however, of whether psycholinguistic results translate in any clear way to the models created by theorists. For example, if we conclude that at least some words must be decomposed in the process of lexical access, can we tell anything from the experimental results about what those rules should look like? What does morphological computation look like, and when does it occur? More than two decades ago Badecker and Caramazza (1989) noted the lack of convergence between psycholinguistic and theoretical study of morphology, but it seems safe to say that such convergence has not yet taken place on any large scale.

6.7 CONVERGENCE IN METHODOLOGY

There are signs, however, that this situation may be changing. The seeds of convergence seem to be appearing in work that is moving away from the practice of using data generated by intuitions to build models loosely based on current syntactic or phonological frameworks. The new convergence acknowledges that corpus-based data and psycholinguistic experimentation are intimately related: the use of vast corpora allow us to look at the statistical and probabilistic nature of derivation and this in turn provides input to experimental design (Hay 2002, Hay and Plag 2004, Plag and Baayen 2009). Here I will just mention one case of the sort of synergy between corpus-based, statistical, and psycholinguistic research that might converge on the theoretical models of the future.

The problem of affix-ordering in English (that is, which affixes can follow which other affixes) has been of keen interest since the early days of generative morphology. Chapter 21 of this volume treats the subject in some detail, so I will only give a brief outline here. An early treatment known as Lexical Morphology and Phonology (LMP) (Siegel 1974, Allen 1978, Kiparsky 1982b) divided the affixes of English into two cohorts, Level 1 and Level 2, each of which was associated with a constellation of phonological rules. Level 2 affixes were predicted to be found outside Level 1 affixes, but not vice versa. Not surprisingly, the predictions of the theory were largely tested on the basis of intuitions. Spencer (1991: 80), for example, cites **hopefulness* and **irrefillable* as words that would be ruled out by the theory. But, as has been pointed out several times, there are complex words that seem to defy the predictions of LMP. In the years since its formulation, there have been many other criticisms of LMP and many attempts to revise it to accord with data (see for example, Giegerich 1999), but it is generally conceded that Level Ordering does not account well for the ordering of affixes.

One reason, pointed out in Fabb (1988), is that it predicts vastly more combinations of affixes to be possible than actually are attested. Basing his case on data that are largely culled from Walker's (1924) rhyming dictionary, Fabb suggests that the ordering of affixes in English is the result of selectional restrictions on particular affixes. Plag (1999) builds on Fabb's results, using data culled from Lehnert (1971) and the *OED*, and suggests that there are restrictions not only on affixes, but also on bases that account for affix ordering.

Hay (2002) adds a new dimension to the debate—significantly one that starts to bridge the gap between purely theoretical models and models of lexical processing and speech perception. She argues that affix ordering is a matter of parsing such that affixes that are more parseable do not occur inside affixes that are less parseable; by parseable she means affixes that (1) give rise to phonotactic transitions that are unlikely to be found morpheme-internally and (2) are found in forms in which the derived word is less frequent than its base. Affix ordering is a gradient matter depending not only on the phonotactics of particular affixes but the internal complexity of specific derived words. In other words, in explaining affix-ordering we need to go beyond just looking at whether affix x occurs outside of affix y, but at the relationship between affix y and its base, for example at the relative frequency of the base to the base+affix y and the phonological relationship between affix x and affix y and between affix y and its base. These sorts of relationships, Hay shows, can affect the possibility of stacking affixes.

This theory of affix ordering has been dubbed Complexity Based Ordering by Plag (2002). What is noteworthy for our purposes is that the sorts of relationships between words that determine Complexity Based Ordering can only be calculated on the basis of a large corpus of data. Hay (2002) makes use of the CELEX database. Subsequent refinements of this sort of approach can be found in Hay and Plag (2004) and Plag and Baayen (2009). Each of these works uses successively larger databases and more sophisticated statistical techniques to gather and analyze data, assess frequency, and model ordering. What the results point to is that affix-ordering is a complex matter involving both selectional restrictions of the sorts long proposed by theorists and processing constraints of

the sort proposed by Hay. This is the sort of convergence that I believe we should look for in the future.

6.8 CONCLUSION

If there is such a thing as a linguistic zeitgeist, I would wager to say that it has been shifting in recent years. The 1950s saw a move from the logical positivist methodology of the American structuralists to the intuition-based methodology of generative linguistics. We are now seeing a further shift: while our goals may still be the mentalist goals of the generativists, we now have the capacity to study derivation in a different way. What I have tried to argue in this chapter is that our increasing ability to make use of vast amounts of natural language data has the potential to profoundly change the way we model the mental lexicon, and specifically the process of forming derived words.

Methodology and theory of course go hand in hand. Both psycholinguistic and corpus data suggest that complex words are not like complex sentences. It seems safe to say that the balance between computation and storage is not an issue in syntax (or at least is less of an issue) and therefore that matters of frequency are more a concern in derivational word formation than in syntax (but see Bresnan and Ford 2010). While we may not be sure what morphological computation is like, it would probably be a good strategy for theorists to assume that it is not a “mini-syntax.” Note that I am not suggesting here that we go back to the Lexicalist Hypothesis of the 1970s which postulated a morphological component of the grammar that was separated from the syntactic and phonological components of the grammar, but that nevertheless proposed rules that looked like either syntactic rewrite rules or rules of phonology. What I am suggesting is that derivational morphology really is different from either syntax or phonology and that theoretical models of the future should start from this premise.

CHAPTER 7

EXPERIMENTAL AND PSYCHOLINGUISTIC APPROACHES

HARALD BAAYEN

THIS chapter provides a critical overview of experimental and computational research on the processing and representation of derived words. It begins with an introductory section addressing methodological issues: The pros and cons of various popular experimental tasks, issues with respect to the selection of materials, as well as the relevance of experimental research for morphological theory. The main section reviews two opposing classes of theories for the organization of the mental lexicon: theories building on the dictionary metaphor, and theories seeking to understand lexical processing without a mental dictionary and without theoretical constructs such as the morpheme.

7.1 METHODOLOGY

7.1.1 Experimental Methods

A wide range of experimental methods is available for probing the processing of complex words. In what follows, some of the more widely used methods are introduced, together with their advantages and disadvantages.

For research on comprehension, the lexical decision task is used widely. Participants are presented with a sequence of stimuli, one at a time, which include both existing words (such as *table*) and non-existing words (such as *flurtle*). They are asked to decide, as quickly and accurately as possible, whether each stimulus is a word or a non-word, by pressing one of two response buttons. Stimuli can be presented in writing on a computer screen (visual lexical decision) or over headphones (auditory lexical decision). The time it takes to execute a response (the response latency or reaction time) as well as

the accuracy of the lexical decision have been found to be highly informative about the processing costs of different kinds of complex words.

The lexical decision task offers several advantages. First of all, it is easy to administer, especially for the study of reading. In recent years, large-scale lexical decision studies have been carried out, collecting reaction times for tens of thousands of words (see, e.g., Balota et al. 1999, 2004, 2007, Ferrand et al. 2010, Keuleers et al. 2010, 2012). At the time of writing, several labs are running experiments using crowd sourcing, with volunteers running lexical decision experiments on remote laptops and smartphones. However, the lexical decision task also has many disadvantages. First, the task requires participants to make a metalinguistic judgment, which is far removed from normal comprehension. Second, words are presented in isolation, whereas in experience words tend to be part of sentences or utterances. As a consequence, lexical decision latencies tend to show only weak correlations with processing measures from the eye-movement record (Kuperman et al. 2013). Third, how the non-words are constructed (see, e.g., Keuleers and Brysbaert 2010) as well as what kind of words are included as fillers in the list of stimuli (see, e.g., Feldman et al. 2009) may substantially affect the results obtained. Finally, reaction times and accuracy scores are uninformative about the time course of lexical processing: Typically, the early stages of lexical processing as revealed by eye-tracking studies can be very different from the later processing stages evaluating lexicality decisions (Miwa et al. 2014), and may even be misleading.

The lexical decision task is often combined with a so-called priming experimental treatment, in which a given target word is preceded by carefully selected other words, the so-called primes. Primes can be words presented at a certain distance earlier in the experimental list (long-distance priming). In the masked priming task (using visual lexical decision), primes are presented for a very short duration (e.g. 60 ms), often preceded by a mask of random letters or hash marks, before the target word is presented, in which case most subjects do not become aware that a prime word was presented. When masked priming is used to study morphological processing, typical priming treatment conditions are an identity condition (*good* priming *good*), a related condition (*goodness* priming *good*), a form condition (*food* priming *good*) and an unrelated condition (*hand* priming *good*). The results one tends to obtain are that responses are fastest in the identity condition, intermediate in the related condition, and slowest in the unrelated (control) condition.

The effect of a masked related prime has been attributed to the prime word partially preparing the way for lexical access for the target word, either by “opening” the lexical entry of the target, or by partially pre-activating the target (Forster 1999). Theorists accepting this interpretation compare the priming effect against the unrelated baseline, in which case a related prime will elicit shorter latencies than the unrelated condition. The reason that a related prime also elicits longer latencies than the identity condition is attributed to a channel capacity problem, with two words having to be processed nearly simultaneously instead of just one word. However, Norris and Kinoshita (2008) argue that in masked priming, the perceptual system cannot properly distinguish between the prime and the target as different perceptual events. As a consequence, the orthographic information of prime and target would blur into one perceptual whole, and the more the

prime differs from the target, the more noise it contributes to the perception of the target, and the longer response latencies become. Norris and Kinoshita (2008) also show that priming effects can be task-specific: present in visual lexical decision, but absent (for the same stimuli) for a same–different task. This implies that the effects of priming need not be an automatic consequence of the structure of the mental lexicon, but arise “online” depending on the demands of the task.

For the study of morphological processing in reading, modern eye-tracking systems offer the possibility of tracing, in considerable detail, and with great accuracy, where the eye lands in a complex word, how often it fixates within that word, and whether the eye will return to the word after having fixated elsewhere (see, e.g., Rayner 1998, Kuperman et al. 2009, 2010). The advantages of using eye-tracking are, in addition to providing detailed insight into the time course of lexical processing, that words can be presented in sentential and/or discourse context, providing experiments with enhanced ecological validity compared to tasks involving lexical decisions. The major disadvantage of eye-tracking is that it is currently impossible to gather data with crowd sourcing. However, this may change in the near future, with the development of user interfaces for smartphones that track where the eye is fixating on the screen.

Experimental research on speech production is much more difficult than research on comprehension. Whereas in comprehension experiments, materials with desired controlled properties can be presented to participants, the challenge in production studies is to get participants to produce the words with the critical properties of interest. In principle, one could present words in writing and ask participants to read them out loud, but this has the serious drawback that results conflate an initial comprehension process with a subsequent production process. Three tasks have been widely used in research on morphological processing in speech production: picture naming, implicit priming, and the picture–word interference task.

For studying speech production from initial conceptualization to final articulation, the picture naming task is a good choice. In this task, participants are presented with line drawings or photographs, and are asked to say out loud as quickly and accurately as possible what the picture denotes. In this task, the input is non-linguistic, and hence the response variables—naming latency and accuracy—gauge the costs of preparing for speech without contamination from linguistically mediated comprehension. The task has two disadvantages, however: Only picturable nouns, verbs, and adjectives can be presented, and the temporal information obtained is restricted to the onset of articulation.

The picture–word interference task seeks to obtain information about the time course of lexical processing by combining picture naming with the presentation of distractors, typically words presented visually or auditorily with the picture. The critical manipulation here is the amount of time between the presentation of the distractor (e.g. *lace* and the presentation of the picture (of, e.g., a shoelace), the “stimulus–onset asynchrony” (SOA). Distractors can be phonologically, semantically, or morphologically similar to the target, and different kinds of distractors typically cause maximal interference at slightly different SOAs.

A third task, implicit priming, builds on participants' ability to learn pairs of word associations (e.g. *hand/foot*, *beach/sea*, *dog/cat*), where the idea is to use the associate (e.g. *hand*) to elicit the target (e.g. *foot*). For training, pairs of words are selected such that the target words either share some critical property (e.g. they might all begin with the same phoneme, the homogeneous condition) or do not share any property (the heterogeneous condition). During testing, only the associates are presented, and participants are requested to say the corresponding targets. Response measures are reaction time and accuracy. Implicit priming has as advantage that targets are no longer restricted to being picturable, but it has as its disadvantage that participants have to perform a rather strange associative memory task with little ecological validity.

Electroencephalography (EEG, the recording of electrical activity on the scalp) and magnetoencephalography (MEG, the recording of magnetic fields produced by electrical currents in the brain), and functional magnetic resonance imaging (fMRI, the mapping of brain areas with increased blood flow in response to experimental events) have made it possible to investigate the details of the time course of lexical processing, as well as the regions in the brain that subserve these processes. These experimental techniques can be combined with behavioral tasks (lexical decision with or without priming, see, e.g., Morris et al. (2007) as well as with eye-tracking (Dambacher and Kliegl 2007) and picture naming Jescheniak and Levelt (1994)). Electroencephalography and MEG come with a high temporal resolution, whereas fMRI comes with high-quality information on localization. While these techniques have the obvious potential of providing detailed information about the temporal and spatial reflexes of linguistic processing in the brain, they also come with disadvantages.

One disadvantage of the neuroimaging approaches is methodological in nature. Especially in the case of fMRI, there are so many choice points in the course of data analysis that for any given study it can be entirely unclear whether results published as "significant" are actually obtained thanks to a fishing expedition in analytical parameter space (Haller and Bartsch 2009, Vul et al. 2009, Carp 2012, Eklund et al. 2012). For the analysis of EEG data, a serious disadvantage in the past has been that analytical methods were limited to repeated measures analysis of variance applied to selected time intervals in which researchers observed that the waveform for a violation condition diverged from the waveform for the corresponding control condition. As a consequence, the ecological validity of EEG studies using the violation paradigm for studying language processing is questionable: In natural language, ungrammatical or nonsensical words and sentences are extremely rare, whereas in many EEG studies, violations are highly frequent, and little is known about the strategies that subjects adopt to deal with the challenge of dealing with distorted language. Fortunately, recent advances in statistical analysis make it possible to study lexical processing under more natural circumstances (see, e.g., Hauk et al. 2006, Kryuchkova et al. 2012).

Moreover, linguistic research using EEG has focused primarily on a negative inflection in the averaged waveform around 400 ms post stimulus onset (the so-called N400) and a positive inflection around 600 ms post stimulus onset (the so-called P600). The N400 has been linked to semantic violations, and the P600 to syntactic violations. Some

studies have taken this to indicate that words would be understood only after at least 400 ms post stimulus onset. However, words can be read at a rate of 5/second (Rayner 1998), which makes it very unlikely that a word's meaning would become available only during the reading of following words (see, e.g., Rubin and Turano 1992, Segalowitz and Zheng 2009, Kliegl et al. 2012). Another disadvantage coming especially with MEG and fMRI are the high costs associated with these techniques. For studies with no immediate medical benefit, and hence without the generous financial support typical for medical and clinical research, these high costs increase the pressure to publish, which in turn increase the risk of fishing expeditions and post-hoc explanations.

7.1.2 Selection of Materials

As pointed out by Forster (2000), the materials going into many experimental studies are not selected randomly when researchers seek to match stimuli for lexical properties across experimental conditions. Often, researchers use their own knowledge of the language and experimental experience to accept certain items, and reject others. Researchers may have intuitions about what items might work, and which might not. The consequences of non-random stimulus selection is, from a statistical perspective, disastrous. First, results do not generalize beyond the items in the experiment. Second, the risk of replication failure is unnecessarily large. For this reason, the “mega-studies” using thousands and even tens of thousands of words (Spieler and Balota 1998, Balota et al. 1999, 2001, 2007, Lemhoefer et al. 2008, Ferrand et al. 2010, Keuleers et al. 2010, 2012) are extremely important: The risk of adverse effects of undocumented and undocumented selection criteria is much reduced.

A further problem in experimental studies concerns the widespread practice of dichotomizing numeric variables. For instance, high-frequency words might be contrasted with low-frequency words. The problem here is that almost all lexical distributional variables are intercorrelated. Higher frequency words tend to be shorter, they tend to have more lexical neighbors that themselves tend to be more frequent, they tend to be composed of higher-frequency letters and letter pairs, they tend to have more meanings, and to occur in higher-frequency word sequences. Traditional studies depended heavily on analysis of variance, and hence sought to build binary contrasts in frequency (high vs. low frequency) while matching on a subset of other lexical variables. It turns out that from a statistical perspective, this procedure has several severe disadvantages. First, statistical power is reduced (Baayen 2010c): It becomes more difficult to ascertain that an effect is truly there. Second, the materials in an experiment are not a random sample, but a sample with very specific properties that run the risk of being atypical for the population. Third, matching constraints tend to severely reduce the number of items, sometimes to such an extent that matching criteria have to be relaxed in order to be able to run an experiment at all.

This brings us to the linguistic quality of the materials. In some studies, the necessity of having sufficient items in each experimental condition has led to the inclusion

of words that from a linguistic perspective should not have been included. As a case in point, consider the influential study of Rastle et al. (2004). This study contrasts three sets of words: suffixed words such as *worker*, words containing a potential suffix but which are not morphologically complex such as *corner*, and a control group. In this study, *fruitful* is included in the pseudo-suffixed group along with *corner*. The rationale of these authors must have been that *fruitful* does not mean ‘full of fruit.’ However, the authors ignore that *fruitful* in the sense of ‘successful’ contrasts with *fruitless* (‘unsuccessful’) and that we can speak of *the fruits of one’s labors* (see Baayen et al. 2011 for detailed discussion).

A final issue in psychology in general, and unfortunately psycholinguistics is no exception, is a strong publication bias. Experimental studies that failed to find an effect, as well as studies reporting a replication failure, tend not to be published. Even worse, studies replicating earlier work, without adding a newsworthy new finding of their own, are almost impossible to publish. As a consequence, far more significant results have been published than warranted by the alpha levels of the field (see, e.g., Ioannidis and Trikalinos 2007, Ioannidis, 2008, Francis 2013). In other words, unfortunately, a fair proportion of studies report false positives.

7.1.3 Relevance for Morphological Theory

The research goal of theoretical morphology is often conceived of as providing a complete and insightful description of the internal structure of words within and across languages. Such descriptions typically aim for a balance between enumeration and analysis. Furthermore, such descriptions are either neutral with respect to the modality of use (speaking, writing, reading, listening, signing), or they implicitly take a production perspective (especially in generative frameworks). Although experimental research on lexical processing is fraught with methodological difficulties, as outlined above, it nevertheless has the potential of enriching our understanding of how language in general, and morphology in particular, work.

First of all, it is worth noting that the processes of speaking, writing, reading, and listening are very different. For instance, an experienced reader can process 5 words per second, whereas in auditory comprehension, 200 ms typically captures only part of a syllable. In production, one proceeds from the message to a carefully orchestrated sequence of articulatory gestures, whereas in auditory comprehension, the direction reverses, the task now being to map a highly variable speech signal onto meaning. Not only do speech production and auditory comprehension have very different time courses, they are also subject to different constraints. In auditory comprehension, the number of words compatible with the speech input (the competitors in the lexical cohort) reduces as the acoustic signal unfolds, whereas in speech production, the initial processing stages have to deal with semantically-driven competition (e.g. the selection between near-synonyms or between a hypernym and one of its hyponyms).

Furthermore, differences between individual language users may lead to remarkably different use of the possibilities offered by the grammar of “the language.” It is well known that women tend to have slightly superior verbal skills compared to men (Kimura 2000), and this difference extends to morphological processing. Ullman et al. (2002) and Hartshorne and Ullman (2006) observed a frequency effect for regular inflected words in English for women, but not for men. They interpret this finding within the declarative-procedural model of language (Ullman 2004), which basically takes Bloomfield’s (1933) conception of the lexicon and maps it onto neural structures taken to subserve declarative memory (containing the unpredictable) and procedural memory (rule-based processing). Women, but not men, would then have a declarative memory containing some even higher-frequency regular complex word forms. The female/male split, however, is not this absolute. Various studies have replicated stronger frequency effects for regular complex words for females, but these studies also documented weaker, but still significant, frequency effects for males (Tabak et al. 2005, 2010, Balling and Baayen 2008, Lemhoefer et al. 2008).

Differences between speakers may also arise as a consequence of differences in experience with language. Older speakers tend to know more words than younger speakers. The entropy of their vocabularies is greater than that of younger speakers. As a consequence, retrieving words from their mental lexicons requires more time (see, e.g., Baayen 2008: 181). This in turn leads older speakers to rely more heavily on the use of pronouns (see for extensive discussion, Ramscar et al. 2014).

The consequences of experience have recently been well documented for reading. For reasons of experimental convenience, most research on lexical processing is carried out using reading. However, the participants in experimental studies in psycholinguistics tend to be convenience samples from undergraduate students in psychology who are required to participate in experiments for course credit. As a consequence, the results in the published literature are strongly biased in that they describe the performance of predominantly highly-educated students of which a large majority is female (Francis et al. 2001, Sander and Sanders 2006). This has not restrained researchers from drawing far-reaching conclusions about lexical processing in general and the architecture of the language faculty. However, when the population of readers is broadened to include students from vocational tracks, qualitatively very different patterns of reading are observed (see, e.g., Kuperman and Van Dyke 2011, 2013; below, we will return to their work when discussing the balance of storage and computation in the processing of derived words.). Although individual differences are well-studied in (educational) psychology, for many years, many psycholinguists implicitly adopted the model of the ideal native speaker from generative linguistics, and had no interest whatsoever in individual variation. Fortunately, this is now changing.

Beyond task and individual differences, experimental studies are also of interest to morphological theory because they may provide evidence that supports or challenges the adequacy of the cognitive architectures posited by linguistic theories. For example, as mentioned above, it has been argued that the traditional distinction between rules and lexicon can be mapped onto procedural and declarative memory respectively

(Ullman 2004). However, for instance, Ramscar and Gitcho (2007) offer a very different neural theory, contrasting implicit striatal learning for word forms with top-down control processes involving the pre-frontal cortex and the anterior cingulate cortex (see also Ramscar et al. 2013). This alternative approach has far-reaching consequences for theories of the acquisition of morphologically complex words, as shown by Ramscar and Yarlett (2007). Or consider distributed morphology and the separation hypothesis (Halle and Marantz 1993, Beard 1995), according to which morphemes are no longer linguistic signs. Various studies have sought to demonstrate for comprehension that morpheme-forms are necessarily accessed before higher-order structures (Pylkkänen et al. 2004, Solomyak and Marantz 2010). This approach in turn is challenged by studies indicating the involvement of semantics and higher-level knowledge at the earliest stages of lexical processing (Feldman et al. 2009, Kuperman et al. 2010). Discussions such as these have the potential of informing morphological theory about which of several formal architectures are more compatible with the experimental evidence.

Finally, experimental research may shed light on questions that remain unresolved within declarative theories. By way of example, consider phonaesthemes in English (e.g. *glow*, *glimmer*, *glare*, *glisten* where *gl* appears to refer to the emission or reflection of light. Since for any putative phonaestheme, there are many counterexamples (e.g. *glove*, *glue*, *glad*), it is hard to tell from the distributional data alone whether series such as *glow*, *glimmer*, *glare*, *glisten* have processing consequences similar to those of regular morphemes. Experimental studies by Bergen (2004) and Pastizzo and Feldman (2009) indicate, surprisingly, that there are indeed strong similarities with the processing of derived words.

7.2 THE ORGANIZATION OF THE MENTAL LEXICON

Dictionaries for Indo-European languages such as English, French, and Greek are organized by entries which are ordered by a wordform as a key. In order to access the meaning of a word, this form key has to be found first, either by paging through a paper dictionary, or by entering the key into the search slot of an electronic dictionary. Once the relevant entry has been located, its contents become available.

Many theories take the organization of dictionaries as exemplary for the organization of the mental lexicon. Models of reading, for instance, typically assume that comprehension is a two-staged process. First, the word's form entry has to be identified. During this identification process, lexical competition takes place with similar word forms. Once access to the lexical form is completed, this form entry would then provide a pointer to the word's semantic and syntactic properties.

Other theories try to free themselves from the dictionary metaphor. These theories instead make use of the network metaphor. In network models, activation is claimed to

spread from form units to semantic units, crucially without critical mediation by some form of “dictionary” entry or units representing dictionary entries. In what follows, theories building on the dictionary metaphor are introduced first. Most work in psycholinguistics has been carried out within this general approach. Network theories, which seek to model lexical processing without a “mental lexicon,” are discussed next.

7.2.1 Theories Building on the Dictionary Metaphor

7.2.1.1 *Reading*

A central question in research on the processing of complex words is whether morphological structure serves the purpose of facilitating lexical access, that is, the identification of the proper form entry that provides access to semantics. Given a vocabulary of V entries, the complexity of finding an entry is $O(V)$ when a linear search is used, and $O(\log V)$ for a binary search. Whatever algorithm is used, a greater vocabulary implies an increased search problem. Now suppose that the V vocabulary items are grouped into F word families by their first constituent (e.g. all words beginning with *work*, such as *work*, *workable*, *workload*, *workbag*, *workbasket*, *worker*, *working*, *workings*, . . . , would be in one word family), then the initial search complexity is reduced from roughly 50,000 to 15,000 (counts based on the celex lexical database, Baayen et al. 1995). Since word families tend to be small (in English, the median family size is 2, with a range of 1 to 187 for content words, and a range of 1 to 433 if prefixes are included), it has been proposed that finding the form entry can be speeded up by breaking down the search problem into an initial word family based search, followed by a second search in the much reduced search space of the word family itself. Early models of lexical access worked out this idea under the assumption of lexical searches being linear searches through frequency-ordered lists of entries (see, e.g., Knuth 1973, Taft and Forster 1975, 1976).

In this approach, two lexical-distributional measures have played an important role as diagnostic litmus tests. If the initial search takes place on the basis of the first constituent, then this search should be completed earlier the more frequent this constituent is. In linear search models, the assumption is that the constituents are ordered by frequency, with the highest frequency forms first in the list. Hence for higher frequency words, the number of search steps required is shorter, which predicts shorter processing times. In interactive activation models (McClelland and Rumelhart 1981)—network models in which interconnected nodes excite or inhibit each other—higher-frequency words are assigned higher resting activation levels, allowing these words to be stronger competitors, which enables them to suppress similar words more quickly. Thus, the frequency of the first constituent becomes a diagnostic for lexical access taking place through morphological decomposition: The visual input is parsed into its constituents, of which the first is used as a pointer to its word family.

The second diagnostic is the frequency of the complex word itself. The more frequent the complex word is, the faster it should be accessible. In serial search models, this second frequency effect is accounted for by ordering the entries in the word families by

frequency. In interactive activation models, nodes for constituents pass on activation to nodes representing whole words. These whole-word units of higher-frequency words are assigned higher resting activation levels, which allows them to reach threshold activation level more quickly than low-frequency complex words (see, e.g., Taft 1991, 1994). Thus, constituent frequency effects are attributed to early morphological decomposition, whereas whole-word frequency effects are attributed to subsequent recombination (Taft 2004), to look-up within word families, or to whole-word nodes that receive their activation from constituent nodes lower down in the interactive activation hierarchy.

An important property of this general approach is that the morphological parsing process is assumed to be blind and automatic. Whenever a potential base is encountered, it is assumed to be parsed out, and to serve as a key to a word family. Taft and Forster (1975) and Taft (1981) argue that when prefixed words are read, the prefix is stripped off, and access proceeds on the basis of the stem. Since prefixes tend to have prefix families that are larger than the word families of their stems, prefix stripping is supposed to provide computational efficiency (Knuth 1973) (actual corpus-based estimates suggest otherwise, however, see Schreuder and Baayen 1994). Under blind decomposition, prefixes are also argued to be stripped off in unprefixed words such as *precipice* and *unique*. Taft and Forster (1975) and Taft (1981) provide experimental evidence suggesting that for such pseudo-prefixed words, prefix stripping comes with a processing cost, as *cipice* and *ique* are not valid access keys to word families.

In the more recent literature, the role of blind obligatory decomposition has focused on pseudo-suffixed words such as *corner*, where a parse into a stem *corn* and a suffix *-er* is possible, but misleading. A large number of studies using the implicit priming paradigm have argued that *corn* is parsed out of *corner* just as *work* is parsed out of *worker*, as the amount of facilitation (with respect to an unrelated baseline) obtained by presenting *corner* and *worker* as primes for *corn* and *work* respectively was found to be equivalent; see Rastle et al. (2004), Rastle and Davis (2008), Kazanina (2011), Lavric et al. (2007, 2012), and also Devlin et al. (2004), Solomyak and Marantz (2010), Lewis et al. (2011). Studies using overt priming have reported similar results (see, e.g., Smolka et al. 2009). The evidence for obligatory decomposition is not unequivocal, however, as other studies reported evidence for truly affixed words having a processing advantage over pseudo-affixed words (Diependaele et al. 2005, Christianson et al. 2005, Morris et al. 2007, Feldman et al. 2009, Dunabeitia et al. 2011). Furthermore, an fMRI study by Bozic et al. (2007) suggests that brain regions with a reduced Bold response (i.e. regions with reduced oxygenation compared to an unrelated control condition) for form relations (*corner/corn*) are distinct from brain regions showing a reduced Bold response for semantic relations (*notion/idea*). Interestingly, both areas show a reduced Bold response for morphologically complex words (*boldly/bold*). Although the Bold response is slow, and does not provide information about the earliest stages of lexical processing, this pattern of results suggests that if indeed there is early morpho-orthographic parsing, it does not have long-lasting effects on lexical processing.

There are various reasons for the lack of consistency in the literature on the possible role of blind obligatory decomposition. First, from a linguistic perspective, stimuli

selected as pseudo-affixed are semantically heterogeneous (e.g. *fruitful*, as mentioned above, see Baayen et al. 2011, for detailed discussion). Second, the nature of the filler materials can influence the strategies used by subjects to meet the task requirements (Feldman et al. 2009). Third, whether pseudo-affixed words side with unrelated controls or with truly affixed words may possibly vary depending on participants' spelling skills and vocabulary size (Andrews and Lo 2013).

Perhaps the most important problem with obligatory decomposition is why it would take place. The central functions of morphology are semantic (Lieber 2004) and syntactic (Kastovsky 1986), rather than to provide some efficient hash code for lexical access in reading. Prefix-stripping might seem advantageous, but when the distributional properties of languages such as English and Dutch are considered carefully, the disadvantages of obligatory decomposition outweigh the advantages. A further complication is that the majority of derived words have idiosyncratic shades of meaning. For instance, a *worker* is, according to the online *Merriam Websters*, "one that works especially at manual or industrial labor or with a particular material," "a member of the working class," or "any of the sexually underdeveloped and usually sterile members of a colony of social ants, bees, wasps, or termites that perform most of the labor and protective duties of the colony." Obligatory decomposition of *worker* into *work* and *-er*, combined with subsequent processes of compositional semantics, will never be able to reconstruct the conventionalized meanings of *worker* (see, e.g., Baayen et al. 2013, Pham and Baayen 2013). The only way in which *work* and *-er* can be constructed to work properly is to construe them as hash codes for table look-up to the abovementioned meanings. Unfortunately, hash coding is an engineering solution that fails to predict semantic effects in lexical processing. Furthermore, there are non-morphological engineering solutions that perform better (e.g., letter trees, see Sproat 1992).

Instead of assuming obligatory morpho-orthographic decomposition, Giraudo and Grainger (2001, 2003) have argued that all words have an orthographic access representation (the dictionary key to meaning) that is activated from the visual input. Once such an access representation reaches threshold activation (suppressing its competitors), the corresponding meaning or meanings become available. Effects of morphological constituents observed across a wide range of studies using primed and unprimed lexical decision as well as eye-tracking (see, e.g., Burani and Caramazza 1987, Laudanna et al. 1994, Laudanna and Burani 1995, Burani et al. 1997, Feldman 2000, Bertram et al. 2000b, Kuperman et al. 2010, Miwa et al. 2014) are explained in this theory as arising due to post-access processes. Once the meaning of *goodness* is understood as 'the quality or state of being good' (ignoring for ease of exposition its use as an interjection expressing mild surprise), activation would fan out to the meaning *good* and from there to the corresponding access representation. In other words, according to this supralexical theory, constituent effects arise as a consequence of having accessed a word's meaning, instead of reflecting mediation by the constituents of the access to meaning.

Parallel dual route models present a hybrid of the obligatory decomposition theories and the supralexical theories. These models assume that form representations exist for both whole words and constituents. Two processes run in parallel and independently,

a direct route and a parsing route. The first route to provide access to a word's meaning is hypothesized to determine behavioral measures such as response latencies, as well as fixation durations. The parsing route operates on the access representations of the constituents, and attempts a combinatorial interpretation. The direct route operates on the word's access representation, and makes use of a pointer from this access representation to the word's semantics (Burani and Caramazza 1987, Caramazza et al. 1988, Frauenfelder and Schreuder 1992, Schreuder and Baayen 1995). There are several reasons for positing a direct route in addition to a parsing route. First, a dual route system is more robust and more efficient than a single-route system (see, e.g., Baayen et al. 1997a). Second, the presence of whole-word access representations provides some protection against the many possible competing morphological parse trees that arise in the morpheme-driven route (Baayen and Schreuder 2000). Third, dual route models are supported by eye-tracking studies indicating that first fixation durations are co-determined not only by constituent frequencies but also by whole-word frequencies (Pollatsek et al. 2000, Kuperman et al. 2008, Kuperman et al. 2009, Miwa et al. 2014).

A serious problem for dual route models is that the two routes appear not to work independently. Across several experiments, a tug of war between constituent measures and whole-word frequency has been observed. For instance, Kuperman et al. (2008) observed the effect of compound frequency to be strongest for the modifiers and heads with smaller word families, and Kuperman et al. (2009) observed a similar interaction of compound frequency by modifier frequency. For Dutch derived words, the effect of whole-word frequency was modulated by suffix length (Kuperman et al. 2010). Such interactions are also present in lexical decision (Baayen et al. 2007). An attempt to address these kinds of interactions using probabilities defined over morphemes, complex words, and word families can be found in Kuperman et al. (2008). However, as more refined statistical methods that have become available for addressing numerical interactions in experimental data (Wood 2006, Baayen et al. 2010) typically show even more complex patterns, they challenge explanations invoking a morphemic probability calculus. A further complication is that the tug of war between constituent and whole-word properties has been found to vary systematically between readers as a function of education level and reading skill. Skilled readers revealed strong lexical competition between whole words (*worker*) and base words (*work*), while poor readers received a processing advantage from higher-frequency base words (Kuperman and Van Dyke 2011).

Kuperman et al. (2010) provide an example for Dutch of the complexities that arise when reading suffixed words in sentential context. Focusing on words with a single fixation, the duration of this fixation is co-determined by how far into the word the eye lands (landing positions that are too early or too late induce longer durations), by the length of the preceding word (the longer the preceding word, the longer the duration), and by the plausibility of the word in the sentence (the more plausible, the shorter the duration). Single fixation durations are also shorter for more frequent suffixed words. However, this effect decreases with increasing length of the suffix, and is totally absent for the longest suffixes (length 5). In other words, when the suffix has substantial support from the visual input, because it is long, the effect of word frequency disappears.

Furthermore, for words with longer suffixes, processing costs increase with increasing imbalance of the morphological family sizes of base and suffix, as gauged with the Kullback–Leibler divergence (Milin et al. 2009a, 2009b). The more the family size of the base and the family size of the suffix are similar, the shorter the fixation durations are. Dual-route theory does not provide predictions of this complexity, and it is unclear how it could be modified to do so.

The discussion thus far has addressed research investigating how “dictionary entries” are accessed from the visual input, with special attention to the role of the form representations for the whole word and its constituents. However, the paradigmatic relations between complex words within word families has also been found to have consequences for lexical processing. Here, it is useful to make a distinction between morphological family size, the type count of words in a word’s morphological family (defined as the set of all words sharing that word as a constituent), and morphological family frequency (the summed frequencies of all complex words in the word family). Various studies on Dutch (Schreuder and Baayen 1997, Bertram et al. 2000a, De Jong et al. 2000) indicate that the predictor relevant for predicting visual lexical decision latencies is the family size measure and not the family frequency measure (but see Ford et al. 2010). The processing advantage of words with large morphological families has been observed also for English (Baayen et al. 1997b, Feldman and Pastizzo 2003, Pykkänen et al. 2004, Baayen et al. 2007), as well as for Finnish and Hebrew (Moscato del Prado Martín et al. 2004, 2005).

The morphological family size effect is usually understood as a consequence of activation spreading in a network of connected dictionary entries from the base of a complex word to its family members. Within the multiple-readout framework of Grainger and Jacobs (1996), the co-activation of many family members provides evidence for a positive lexicality decision. Alternatively, if activation is allowed to resonate within a morphological family, this resonance can significantly boost the activation of the presented word, and hence afford shorter response latencies (De Jong et al. 2003). The family size effect is semantic in nature. This is seen clearly in the results obtained for Hebrew. Moscato del Prado Martín et al. (2005). They studied derived words with homonymic roots such as X-SH-B, which contribute to two semantic families, one involving concepts related to thinking (e.g., *Xa-SHaB* ‘to think,’ *maXSHaBa* ‘a thought,’ *XaSHiBa* ‘thinking,’ and one relating to concepts involving arithmetic (e.g. *XiSHeB* ‘to calculate,’ *XeSHBon* ‘arithmetic,’ *XiSHuB* ‘calculation’). Response latencies turn out to be sensitive to which semantic family (within the root family) a word belongs to. When a derived word from one semantic word family is read, response latencies decrease for increasing family size of that family, whereas response latencies increase for increasing family size of the other, semantically unrelated, root family members. In other words, the effect is sensitive not just to the presence of a shared consonantal root, but to the semantic fields supported by a given root.

In the reading research, there are several other lines of research relevant for the processing of derived words. To continue with Semitic, masked priming studies on Hebrew and Arabic show facilitation for prime–target pairs sharing the consonantal root,

compared to unrelated controls (Frost et al. 1997, 2000a, 2000b, Deutsch et al. 1998, Bentin and Frost 2001, Boudelaa and Marslen-Wilson 2001, 2004, Frost et al. 2005, Boudelaa et al. 2009). For Hebrew (Deutsch et al. 1998), but not for Arabic (Boudelaa and Marslen-Wilson 2004), primes sharing the vowel pattern but not the consonantal root also facilitate responses. Research on lexical processing in Semitic interprets the experimental results on the processing of roots and vowel patterns as straightforward evidence for the cognitive reality of morphemes. Surprisingly, experimental psychologists seem to be unaware of alternative linguistic analyses of non-concatenative morphology such as proposed by Ussishkin (2005, 2006).

One of the striking properties of reading is that it is difficult, for words such as *answer*, to detect misspellings consisting of letter transpositions. Within monomorphemic words, masked primes with a letter transposition have been found to be almost as effective as identity primes (Perea and Lupker, 2004). For complex words, letter transpositions within constituents are also nearly harmless, but transpositions at the morpheme boundary (e.g. *db* in *sandbank*) have been found to be disruptive in some studies (Christianson et al. 2005, Dunabeitia et al. 2007, Lemhöfer et al. 2011), but not in others (Perea and Carreiras 2006, Rueckl and Rimzhim 2011). It remains at present unclear to what extent manipulation of the boundary bigram in complex words can serve as a diagnostic for morphological processing.

A final question in reading research asks whether morphology is more than just the coincidence of shared form and shared meaning? Feldman (2000) addresses this question and shows that in primed lexical decision, morphological effects exceeded the individual effects of semantic similarity and of form similarity. Her conclusions find support in more recent neuroimaging studies (Bick et al. 2008, Boudelaa et al. 2009) which suggest that there are brain areas that are involved only when morphologically complex words are processed, and not for words related in only form or only meaning. It should be kept in mind, however, that similar conclusions are reached (Pastizzo and Feldman 2009) for word pairs such as *boat-float* (semantically related and related in form), *swim-float* (only semantically related) and *coat-float* (only related in form). In other words, the sharing of form and meaning seems to be important, and not whether or not this sharing is brought about by means of affixation.

7.2.1.2 *Listening*

When reading, morphological processing of derived words is influenced by factors such as where the eye lands, how much of the suffix is visible, and the lengths and frequencies of the constituents. When listening, instead of information about large chunks of words becoming available simultaneously, the speech signal unfolds slowly over time. As a consequence of this slow temporal unfolding, the set of words compatible with the input is winnowed down as more information becomes available. For instance, after having heard just the first two segments of *houseful*, the word *hound* is still a viable continuation, but after having heard the third segment, only words beginning with *house*, including *house* itself, remain in the cohort of lexical competitors.

Marslen-Wilson (1996) proposed a definition of competitors based on a morphological breakdown of the lexicon. Only uninflected monomorphemic words were included, and derived prefixed words. On the basis of this set of potential lexical competitors, he defined a word's uniqueness point as the point in the speech signal at which all of a word's competitors have become incompatible with the speech input. Other things being equal, a word with a uniqueness point earlier in the word tends to be recognized more quickly.

Suffixed and compound words are not considered in cohort theory for two reasons. First, their inclusion would give rise to most words becoming unique after word offset, which would be self-defeating. Second, in the 1980s, compositional theories were dominant, and no experimental evidence was available on, for instance, the importance of whole-word frequency as an independent predictor of the processing complexity of regular complex words in auditory comprehension.

Nevertheless, standard cohort has been shown to be too restrictive to be revealing about morphological processing. Following up on work by Wurm (1997), Wurm and Ross (2001), Wurm and Aycok (2003), and Balling and Baayen (2008), Balling and Baayen (2012) defined two uniqueness points. The initial uniqueness point (UP) is reached when morphologically unrelated competitors are no longer compatible with the speech input. The complex uniqueness point (CUP) is reached when morphologically related competitors drop from the cohort. Experiments using the auditory lexical decision task show that both uniqueness points predict shorter response latencies for words with earlier UP and CUP. Balling and Baayen (2012) proposed to understand uniqueness point effects as reflecting changes in surprisal (approximately, amount of information), in parallel to the way changes in surprisal predict processing costs in syntax (Levy 2008). By the time a uniqueness point has been reached, most of the cognitive costs associated with weeding out unrelated competitors have accrued. As a consequence, subsequent processing can proceed more quickly.

One of the consequences of the distribution of the language signal in time (spoken) rather than space (written) is that morphological family size effects are less robust (see Balling and Baayen (2012) for detailed discussion). Family size counts are insensitive to position: Complex words are counted irrespective of whether the targeted base word occurs in initial position. However, for auditory comprehension, order does matter. Although the family size count of *house* includes words such as *rehouse* and *roadhouse*, these family members have long dropped out of the cohort when listening to *house* itself. Balling's complex uniqueness point is therefore a more useful construct for gauging paradigmatic morphological structure, as the cohort of competitors that is active between the UP and the CUP consists of morphologically related words. Nevertheless, a family size effect was detected in the EEG signal, starting around 150 ms post stimulus onset, elicited in a normal listening task (with isolated words, but without a decision component) by Kryuchkova et al. (2012).

The uniqueness point effects fit well with the whole-word frequency effects observed across many experiments (Meunier and Segui 1999a, b, Baayen et al. 2003, 2007, Balling and Baayen 2012) as well as with research on acoustic reduction (Schuppler et al. 2012),

and point to a rich lexicon with semantic representations not only for monomorphemic words, but also for complex words. Interestingly, the computational model for auditory comprehension of Norris and McQueen (2008) is also based on such a rich lexicon. According to this model, the probabilities of form representations for both simple and complex words undergo continuous Bayesian updating as the speech signal unfolds. No specifically morphological processes are involved during listening. It is unclear, however, how this model would handle the understanding of novel complex words that are not in its lexicon. For cohort theory, it is convenient to think of the speech signal as a series of discrete segments. However useful for formulating theoretical constructs such as uniqueness points, the actual speech signal is much richer. For understanding morphological processing in auditory comprehension, the richness of fine phonetic detail is an important factor to be taken into account.

A comparison of the orthographic forms of *work* and *worker* would suggest that information about morphological complexity is carried exclusively by the suffix. However, in speech, the prosodic cues of the stem change when the suffix is added. For instance, syllable structure changes, and the stem becomes shorter. As a consequence, listeners can already anticipate upcoming morphological structure while listening to the stem (Kemps et al. 2005a, 2005b). Furthermore, in colloquial speech, complex words are often produced in highly reduced form. For instance, the Dutch adverb *eigenlijk*, [ɛɪxənɫək], is often reduced to single-syllable [ɛxk] (Ernestus 2000, Keune et al. 2005). Out of context, such strong reductions are difficult if not impossible to understand (Ernestus et al. 2002), whereas successful understanding of a strongly reduced form appears to come with the percept of a much richer, more canonical, phonological form (Kemps et al. 2004). In some cases, the fine phonetic detail of the reduced form still reflects its polysyllabic origin (Niebuhr and Kohler 2011), which may help guide the listener to the appropriate meaning.

Acoustic reductions of (complex) words, like the paradigmatic effects discussed in the previous section, challenge the usefulness of the dictionary metaphor for lexical processing. It is, of course, possible to enrich the lexicon with separate auditory access representations for reduced words, but such a move does not help explain why without context [ɛxk] does not activate any semantics. Given frequency effects observed for regular (non-idiomatic) word sequences (Bannard and Matthews 2008, Arnon and Snider 2010; Tremblay and Baayen 2010, Tremblay et al. 2011), it might be argued that reduced forms are part of multiword templates, and that the mental dictionary should be broadened to a repository of both words and phrases. However, this could lead to hundreds of millions of additional entries for canonical *n*-word combinations (with *n* < 5) alone. A more dynamic approach with context-sensitive anticipation of the acoustic consequences of admissible articulatory shortcuts would not have this disadvantage (cf. Baayen et al. 2012).

7.2.1.3 *Speaking*

As holds for theories of language comprehension, models of speech production are also heavily influenced by the dictionary metaphor, with a reversal in direction being a

major change: The speaker, with some communicative goal in mind, has to find the right words to express herself. Current models posit nodes (representations) for word meanings, and link these meanings up to nodes for word forms, which in turn may link up to nodes for syllables and/or phonemes. The two main models in the literature differ in how activation is passed on from one node to the other. The interactive activation model of Dell (1986) posits both top-down and bottom-up links, which causes lexical processing to become highly interactive. In the weaver model of Levelt et al. (1999), connections are strictly top-down, from higher conceptual levels to lower levels of word forms and segments. In both models, various rules and checking mechanisms ensure that at the different levels the proper nodes are selected. What both models also have in common is that morphologically complex words are assumed to be constructed from their constituent morphemes.

In the model of Levelt et al. (1999), morphemes are not conceptualized as the smallest meaning-bearing unit, but as formal planning units. The complete separation of form and meaning, which fits well with the separation hypothesis of Beard (1977, 1981a, 1995) as well as with distributed morphology (Halle and Marantz 1993), is motivated by a series of experiments using the implicit priming paradigm. These experiments suggest that the semantic compositionality of a complex word is irrelevant for speech production: semantically transparent words such as *input* and semantically opaque words such as *invoice* exhibited a priming effect of similar magnitude that was much larger than for monomorphemic controls such as *insect* (Roelofs and Baayen 2002). Using a long-distance priming paradigm with picture naming, Koester and Schiller (2011) reached the same conclusion, as did Lüttmann et al. (2011) using picture-word interference experiments. For Hebrew, however, Deutsch and Meir (2011) reported effects of morphology that did not reduce to the joint effects of form and semantic similarity.

Consistent with a strictly decompositional approach to speech production, Roelofs (1997) obtained a base frequency effect for particle verbs. On the other hand, experimental evidence for whole-word frequency effects in speech production is mixed. A picture naming study on plural inflection in Dutch failed to find a whole-word frequency effect (Baayen et al. 2008) which was well attested for similar word materials in the same language for both reading and listening (Baayen et al. 1997a, 2003). Although Bien et al. (2005) observed a U-shaped frequency effect for compounds in a position-response association task, Bien et al. (2011) failed to find whole-word frequency effects for inflected and derived words. However, Tabak et al. (2010) observed effects of form frequency across several picture naming experiments with inflected verbs, and Janssen et al. (2008) found strong support for a whole-word frequency effect for compounds in picture naming in both English and Chinese. As pointed out by Janssen et al. (2008), it is quite possible that more natural tasks such as picture naming are better suited for the detection of whole-word frequency effects than associative memory tasks such as implicit priming or position-response association.

Further challenges to strictly decompositional models of speech production come from two sources. First, for inflected words, the entropy of the inflectional paradigm has been found to predict response latencies in both the picture naming (Baayen et al.

2008, Tabak et al. 2010) and the positional response association task (Bien et al. 2011). The inflectional entropy measure can be thought of as estimating the difficulty of choosing between different inflectional variants. The greater the inflectional entropy, the greater this difficulty is, and the longer response latencies become. Inflectional entropy effects show, albeit indirectly, that the speech production process is sensitive to the relative probabilities of inflected wordforms. Since strictly decompositional models have no representations for inflected wordforms, they cannot predict relative entropy effects.

A second challenge for strictly decompositional models of speech production comes from analyses of the speech signal. Whereas experimental studies in speech production typically work with response latencies, or with the consequences of priming manipulations on the electrophysiological response of the brain (Koester and Schiller 2011), the phonetic record of what speakers have actually said is also highly informative (see, e.g., Gahl 2008). Using a large speech corpus, Pluymaekers et al. (2005b) were able to show that the acoustic durations of prefixes and suffixes and/or the durations of segments in these prefixes and suffixes may be co-determined by the frequency of the derived words in which they occur. Pluymaekers et al. (2005a) showed, furthermore, that the acoustic realization of a suffix is co-determined by contextual factors such as the number of times the word was used in the preceding discourse, as well as its predictability from the preceding and following word. In addition, Tremblay and Tucker (2011) observed that the frequency with which combinations of four words occur co-determines acoustic duration. In the model of Levelt et al. (1999), which posits that the segments selected by morphemes are first bundled up into syllables, and which takes these syllable units to drive articulation, it is difficult to see how the frequency of a higher-order unit of a derived word (represented in the model only at higher conceptual and syntactic levels, but not at the wordform level), and contextual probabilities, might affect the articulatory execution of an affix. More in general, the *WEAVER* model of speech production is challenged by the accumulating evidence that a word's similarity neighborhood co-determines speech production (Vitevitch 2002, Munson and Solomon 2004, Scarborough 2004).

A similar challenge comes from work on relative frequency. Hay (2003) distinguished between derived forms which are more frequent than their base words (e.g. *illegible*, *swiftly*) and those derived words for which the base is more frequent (e.g. *illiberal*, *softly*). Hay observed more t-deletion in English for derived words with a large relative frequency (*swiftly*, derived frequency > base frequency) than for words with a small relative frequency (*softly*, derived frequency < base frequency). A similar effect of relative frequency was reported for Dutch by Schuppler et al. (2012), but not in a reanalysis by Hanique and Ernestus (2012). If the effect of relative frequency turns out to receive more experimental support, it challenges full decomposition production models. By denying a role to whole-word representations for complex words, it becomes impossible to predict segment reduction from form frequency.

Thus far, we have considered the production of speech. Some results are also available for the production of writing. A large series of studies on typing in German, reviewed in Weingarten et al. (2004), investigated inter-keystroke intervals. For letter pairs spanning a morpheme boundary but not a syllable boundary, inter-keystroke intervals

did not differ for non-morphological controls. However, when morpheme and syllable boundaries coincide, inter-keystroke intervals were found to be longer compared to controls with only a syllable boundary. Weingarten and colleagues also observed an effect of whole-word frequency, independently of base frequency. Kandel et al. (2012) compared, for handwriting, interletter pauses at the morpheme boundary for prefixed words and suffixed words in French, and compared them with pseudo-affixed controls. They only found a difference for suffixed words, from which they conclude that only suffixed words would be processed compositionally.

7.2.2 Lexical Processing without a Mental Lexicon

The theories and models reviewed thus far build on three important assumptions. First, they all accept without question that there are discrete lexical units for morphemes. Questions raised about the validity and usefulness of the morpheme as a theoretical construct, as raised by Matthews (1974), Uhlenbeck (1978), Anderson (1992), Stump (2001), and Blevins (2003), have not entered into the awareness of most of the psycholinguistics community.

Second, the models formulated in this framework, irrespective of whether developed only as blueprints or computationally implemented, are declarative models that systematize a large body of knowledge, but, importantly, that do not learn. Irrespective of whether a dictionary theory works with interactive activation or with just a unidirectional flow of activation, the algorithms are designed to work in exactly the same way for a given word, irrespective of how many times that word (and other words in its context) have been encountered.

Third, these models work with a highly idealized and simplified view of the relation between form and meaning. Here, several issues come into play. First, from a linguistic perspective, it makes sense to distinguish between the skeleton and body of a word's meaning (Lieber 2004), where the skeleton denotes the language structural scaffolding that supports the body, the rich encyclopedic knowledge that is part of a word's meaning. Thus, returning to the above example of *worker*, the skeleton is (simplified) "a subject noun derived from the verb *to work*," whereas part of the body is that the word denotes a particular kind of bee. It is important to realize that theories of lexical processing have to explain, for instance, how a listener comes to a proper understanding of a sentence such as "In the warm afternoon sun, we could see many workers collecting honey". Since no rule can reconstruct the meaning "bee" from *work* and *-er*, compositional theories of comprehension can only provide access to the skeleton, but not to the body. Similarly, compositional theories of production cannot account for the longer acoustic duration of *worker* in the low-frequency sense of "honey bee" compared to the high-frequency sense of "participant on the industrial labor market" (cf. Gahl 2008). Furthermore, it is not the case that in comprehension, the skeleton is accessed first, subsequently to be enriched with its body. Evidence is accumulating that rich information about the body plays an important role already during the earliest stages of

comprehension (see Elman 2009, and references cited there). Finally, words do not have or carry meanings (Ramscar and Baayen 2013)—it is only thanks to the context in which a word occurs that they come into their own (recall, for instance, that strongly reduced derived words, for which we do not have orthographic awareness, are not interpretable out of context, see Ernestus et al. 2002, Kems et al. 2004).

Two kinds of approaches have been pursued for understanding lexical processing without mediation by form entries for words or morphemes. Both take learning very seriously. Distributed connectionist models (McClelland and Elman 1986, Norris 1994, Joanisse and Seidenberg 1999, Seidenberg and Gonnerman 2000, Bird et al. 2003, Moscoso del Prado Martín et al. 2003, Moscoso del Prado Martín 2003, Harm and Seidenberg 2004) seek to explain morphological effects in the experimental literature as an emergent property of a processing architecture with three interacting banks of units: a bank of orthographic feature units, a bank of phonological feature units, and a bank of semantic feature units. In the TRIANGLE model (Harm and Seidenberg 2004), each of these banks of units is connected through intervening banks of hidden units. These hidden units serve a dual purpose: They allow for compression of statistical regularities between form and meaning, and as a consequence similarities in patterns of activations over hidden units can come to resemble generalizations over the input space. Seidenberg and Gonnerman (2000), Plaut and Gonnerman (2000), and Gonnerman and Anderson (2001) made use of distributed connectionist models to explain processing advantages in priming studies of derived words (*boldly*–*bold*) vis-à-vis orthographic (*corner*–*corn*) and semantic (*idea*–*notion*) controls as arising due to the convergence of form and meaning.

Although distributed connectionist models are learning models, the learning algorithm used, back-propagation, has been criticized for being psychologically and neurobiologically implausible (Crick 1989, Murre et al. 1992, O'Reilly 1998, 2001). Furthermore, designing such models involves choices about the number of banks of hidden units, the numbers of units in the different banks, and the featural representations chosen for orthography, phonology, and semantics. A further criticism of these kinds of models is that the behavior of any given model requires detailed statistical analysis of the banks of hidden units.

An alternative to distributed connectionist models is the naïve discrimination learning (NDL) model (Baayen et al. 2011). The network structure of an NDL model is extremely simple: the nodes in a first layer of cues are linked up to the nodes in a second layer of outcomes. There are no hidden layers, and both cues and outcomes are straightforward symbolic representations (e.g. cue nodes for letters and letter pairs, and outcome nodes for meanings). The weights on the links from cues to outcomes are estimated from the equilibrium equations of Danks (2003) for the learning equations developed by Wagner and Rescorla (1972). This makes it possible to estimate the connection weights from large corpora with hundreds of millions or even billions of words. Given the weights, the activation of an outcome is obtained by summation over the weights from the cues in the input to that outcome. The activation of a meaning outcome reflects how well that meaning can be learned given the words, their orthographic forms, and their meanings, in the language as sampled by the corpus.

Thus far, NDL modeling studies are available only for reading. The model of Baayen et al. (2011) comprises a network trained on a quarter of the British National Corpus, using letter unigrams and bigrams as input cues, and symbolic representations for meanings (e.g. ‘work’ and ‘agent’ for *worker*) as outcomes. At the level of semantics, the model therefore is a full decomposition model. Interestingly, the model correctly captures whole-word frequency effects, morpheme frequency effects, and family size effects observed with the visual lexical decision task, even though there are no representations for whole words, morphemes, or morphological families in the model’s architecture. Although distributional measures such as word frequency, constituent frequency, and family size, are often interpreted as diagnostic measures for cognitive representations for whole words, for constituents, and links between morphologically related words, the NDL model shows that a very different interpretation is possible, an interpretation in which these effects reflect learnability.

It is important to keep in mind that the NDL model is not a model of the full reading process. To the contrary, the model captures only the very first stage of the reading process, namely the activation of meaning from low-level visual information (represented in the model by letter unigrams and bigrams). Often, more than one fixation will be necessary for understanding a complex word, and the higher-order cognitive processes further guiding interpretation (Yeung et al. 2006, Ramscar and Gitcho 2007) constitute an essential part of reading that is not captured by the NDL (see Baayen et al. 2013).

The NDL model has thus far been applied not only to English but also to Serbian and to Hebrew. The modeling results of Baayen (2012) suggest that skilled reading of Hebrew may not require a non-concatenative decomposition into morphemes as argued for by (McCarthy 1981). Theoretically, the NDL model is much closer to the phonotactic approach of Ussishkin (2005, 2006).

Naïve discrimination learning makes a prediction concerning the role of infrequent phoneme sequences straddling morpheme boundaries that is exactly opposite to what connectionist models as well as symbolic models with morphemic decomposition predict. Hay (2002, 2003) argued that infrequent letter bigrams straddling a morpheme boundary (e.g. *tl* in *swiftly*) would make the complex word more parseable. Likewise, Seidenberg (1987) argued that for a connectionist network to learn word-specific meanings, higher-frequency boundary digraphs are required. As a consequence, words with low-frequency boundary digraphs would depend more on the mappings of form to meaning in the stem and in the affix, thereby giving rise to processing effects (in the network) that in a symbolic framework would be understood as the effects of parsing. By contrast, in naïve discrimination learning, the lower the frequency of a boundary bigram, the better its cue value becomes for the complex word’s own meaning (see Baayen et al. 2013 for a modeling study in which derived words and compounds have their own meaning outcomes). At the beginning of Section 7.2.2, we distinguished, following (Lieber 2004), between the skeleton and the body of a word’s meaning. Specifically, accessing the body, a word’s idiosyncratic senses such as “bee” for *worker*, depends in the NDL model on the boundary bigram. To see this, consider the letter bigram *qa* that appears in the scrabble word *qaïd*. As long as this is the only

word with *qa* known to a reader, the presence of *qa* is a perfect cue to *qaid*. However, the more other words with *qa* exist in a speaker's lexicon (e.g. *qanat*), the less good *qa* is as a cue for *qaid*. In the same way, boundary bigrams that have a low frequency, indicating that they occur in relatively few other words, have a high cue validity for those words. The other side of the same coin is that a low-frequency bigram does not interfere negatively during learning with the activation of the meaning of the base, hence for consonant-initial suffixes, base frequency effects are more likely to be detected (for experimental evidence, see Järvi­kivi and Pyykkönen 2011, Vannest et al. 2011).

The re-evaluation of the functionality of low-frequency digraphs in reading suggested by the NDL approach may also shed light on the comprehension of highly-reduced derived words such as [ɛxk] for [ɛixənlək]. Whenever acoustic reduction results in rare sequences of segments (such as [xk]), these sequences become excellent cues to meaning (see Baayen 2010a for simulations and detailed discussion).

The hypothesis of complexity-based ordering for English derivational suffixes (Hay 2003, Hay and Plag 2004, Plag and Baayen 2009, Baayen 2010b) is also challenged by the NDL approach. Hay's original hypothesis was that suffixes that are more parseable must occur outside of suffixes that are less parseable. It is not self-evident why a parseability constraint of this form should be in force. An alternative description tapping into the same phenomenon is that productivity decreases as one moves from the right edge to the stem (Krott et al. 1999). Since the more productive suffixes tend to be consonant-initial, these suffixes are more likely to create low-frequency boundary diphones/digraphs, which, if the NDL approach is on the right track, would make these words easier to understand. That is, from an onomasiological perspective, consonant-initial suffixes would create words that are more memorable, and hence have higher probabilities of becoming entrenched in the language.

Possibly, stem and suffix allomorphy likewise enhance the discriminability of the different meanings indexed by combinations of stems and affixes (for detailed experimental studies of allomorphy, see Järvi­kivi and Niemi 2002, Järvi­kivi et al. 2006).

Naïve discrimination learning also offers a new perspective on the interactions that often emerge in regression studies of lexical decision and eye-tracking (e.g. Kuperman et al. 2010, Miwa et al. 2014) between measures such as, for instance, whole-word frequency and base frequency. For visual lexical decision, Baayen et al. (2007) observed the strongest effect of whole-word frequency for words with the lowest base frequencies. Conversely, the effect of base frequency was facilitatory for words with low whole-word frequencies, but inhibitory for words with high whole-word frequencies. Within interactive activation frameworks, this pattern of results suggests a tug of war between whole word and base. However, in a learning approach, this tug of war unfolds during the (continuously ongoing) learning process, with cues competing for meanings. Importantly, in real time during reading, there is no actual competition between the meanings of the derived word and its stem, at least during the initial stages of visual processing, as in the NDL model there is just a single forward pass of activation from the orthographic cues to the semantics.

7.3 CONCLUDING REMARKS

There are several critical challenges for research on lexical processing for the coming years. First, the field needs corpora that come closer to actual language experience. Corpora of what people actually say, hear, and read, would be ideal, but since such corpora are prohibitively expensive to develop, corpora using film subtitles are a good approximation (for empirical evidence, see, e.g., Brysbaert and New 2009). The reason for subtitle corpora working well probably is that they approximate more accurately the colloquial use of everyday spoken language.

Second, it will be important to move away from the lexical decision task, as it may tell us more about a metalinguistic judgment task than about actual language processing. However, as psychologists have discovered crowd sourcing and have developed apps for smartphones that can easily harvest millions of lexical decisions, this method will become more instead of less popular in the coming years.

Third, to advance the field, computational implementation is essential. The verbal models of the last 40 years (prelexical decomposition, postlexical decomposition, dual-route models) fail to predict the complex patterns present in the experimental data. If language shares essential properties with complex dynamic systems, which is what the experimental data suggest, then linguistics and psycholinguistics will need to start using the tools and techniques developed in other domains of scientific inquiry for studying complex dynamic systems, and to give up the static dictionary metaphor that still guides many current models of lexical processing.

Fourth, current research on the processing of derived words (and of lexical processing in general) is typologically severely limited, with strong research traditions restricted to selected Indo-European languages, to Semitic, to Finnish, and to Chinese, and Japanese. In all these cases, we are dealing with societies with long traditions of literacy, and with experimental research with a strong bias for the study of reading. It goes without saying that the generality of the results reviewed in this chapter is severely limited by this bias. Finally, the consequences of continued learning throughout the life time and the concomitant accumulation of knowledge (including lexical knowledge) has profound consequences for individual differences in language processing (Ramscar et al. 2014). The field will need to abandon convenience sampling of university students, and to commit to sampling from broader cross-sections of the population if we are to obtain a realistic view of how language really works in our societies.

CHAPTER 8

CONCATENATIVE DERIVATION

LAURIE BAUER

8.1 WHAT IS AN AFFIX?

8.1.1 The Basics

LET us begin with the notion of a lexeme. In the sense in which *wants* and *wanted* represent the same “word” we say that they represent the lexeme WANT. Some lexemes like BRUTALISE have rather more analysable parts than WANT. Nevertheless, at the heart of any lexeme is an irreducible element which links a particular form with a particular lexical meaning. That smallest core of any word, in BRUTALISE the element *brute*, is called the “root” of the word. An affix is a recurrent piece of phonological material, not itself a root, which when found in a word has a relatively consistent effect on the meaning of the word in which it is found. The point about the lexeme is that the form which represents the lexeme may be variable. In inflection, this gives rise to a distinction between word-based morphology and stem-based morphology (see Bloomfield 1933: 225 for the idea and Bauer 2004a for the terms), and the same distinction can be drawn in derivation (*pace* those who restrict the term “stem” to inflectional morphology). Thus, in the English examples in (1), the affixes (in italics) are added to word-forms of English, while in the Italian examples in (2), the affixes are added to something which is not a word-form, and which some term a “stem.” More generally we can say that anything to which an affix attaches is a “base” for that affix. The root is a special kind of base, namely the smallest base in a word.

- (1) command -*ant*
 in- substantial
 king -*dom*
 murder -*er*
 pre- paid

- (2) Italian
 cann-etto ‘reed + affix = reed bed’
 fior-ista ‘flower + affix = florist’
 poll -ame ‘chicken + affix = poultry’
 post -ino ‘post + affix = postman’
 sell -aio ‘saddle + affix = saddler’

(Maiden and Robustelli 2000: 437)

The requirement that an affix must not be a word means that affixes cannot stand alone as utterances (unless they are being mentioned¹): to use a standard piece of terminology they are bound morphs. Where something that is a word is added to another word we no longer have derivation but composition (or compounding, see Chapter 3).

There are cases where this apparently simple distinction becomes difficult to apply. In most cases they can be put down to problems in defining a word, something which it is notoriously difficult to do. To cite just one problem, there are at least three books with the title *Isms and Ologies*. Here we have one affix, *-ism*, being used as word (we will return to the status of *-ology* in Section 8.1.2), contrary to expectations. Although I have no wish to sweep problems under the carpet, I shall ignore such problems here on the grounds that they are extremely rare. A fuller account would need to be able to put these exceptions into the system.

8.1.2 The Problem of Neoclassical Word Formation

There is a set of words in many of the languages of Europe which cause problems for the definitions that have been given above. These are words like *hydrophobia*, *photograph*, *thigmotaxis* where the elements are Greek (sometimes Latin), but the words were formed in the modern period, not in the classical period. Such words are instances, therefore, of neoclassical word formation. Some examples of such formations in English are given in (3). In (3) the first column shows a word made up of two neoclassical elements, columns two and three show the same elements in different positions in the word, and the final column shows the first element attached to a word rather than to another neoclassical element.

(3) morphology	isomorph	logophile	morphogenesis
			mythology
neuralgia		algolagnia	neuropathology
phonograph	telephone	graphology	phonocardiogram
rheotome	otorrhoea	tomography	rheometrical

¹ So in answer to the question, “Can you name an affix of English?” the answer might legitimately be, “dom”; this is mentioning the suffix, not using it.

In (3) we see that many of the Greek elements (but not all) can occur either initially or finally in a word, that there is typically an *-o-* linking the two Greek elements, that there may be a final English suffix (e.g. the *-y* in *morphology* or the *-ical* in *morphological*), and that these Greek elements may also be attached to ordinary words. This last fact, illustrated in the last column in (3), might suggest that these elements are affixes. There are two problems with this suggestion. The first is that, if that is the case, a word like *photograph* is made up of two affixes with no base, contrary to the assumptions that have been made above. The second problem is semantic in nature. The meanings of the elements illustrated in (3) correspond very closely to the meanings of lexemes: *morph* means ‘shape’, *neur* means ‘nerve’, *phon* means ‘sound’ and so on. Affixes typically have far more grammatical-looking meanings (we will return to this point in Section 8.5). This makes the words in (3) look rather more like compounds, and as is the case with native compounds in English, the semantic relationship between the elements is variable: *otorrhoea* is a flow from the ear, but *neuralgia* is pain in the nerves, for instance. If words like *photograph* are compounds, then we must presumably also count words like *morphogenesis* and *mythology* as compounds, albeit compounds that mix Greek with English. And if that is the case, the elements in these words are not affixes. That position will be taken here.

We might also argue about the *-o-* that links the Greek elements: is that an affix or not? We will come back to this question in Section 8.2.1. The *-y* at the end of *morphology* (and so on) is clearly an affix.

There is no assumption made here that other language families will have a phenomenon which corresponds to neoclassical word formation. However, it is perfectly possible that other languages also have word-formation strategies where it is difficult to decide whether one is dealing with affixation or not. Neoclassical word formation in English (and in many other European languages) provides a good example of such a difficulty.

8.1.3 The Paradox of Zero Affixes

Affixes were defined above as being pieces of phonological material, that is, as being forms. In most instances, this is totally uncontroversial. However, there are some examples where a claim is made for an affix which has no form at all.

Consider the set of words in (4).

(4)	Base	Derived word	Meaning of derivative
	liquid	liqu-efy	‘to become liquid’
	red	red(d)-en	‘to become red’
	white	white-n	‘to become white’
	yellow	yellow	‘to become yellow’

The first three words have an affix, but the verb *yellow* has no overt affix at all. Some scholars have accordingly suggested that, to draw the parallel between the various cases,

an affix should be postulated on the end of the verb *yellow*, but one which has zero form. The verb would then be written as *yellow-Ø*. Such an affix is called a zero-affix.

This leads to a paradox: a form cannot have no form. The analysis with zero-affixes, called zero-derivation, is correspondingly controversial, though widespread (for discussion and a pro-zero view, see Kastovsky 2005; for a criticism of zeroes, see Štekauer 1996).

8.1.4 Unique Affixes

Some affixes do not seem to recur. There are relatively few examples in English, but a well-discussed one is *-ric*, which occurs only in *bishopric* (and the closely related *archbishopric*, where it has the same base). In a case like this, it is impossible to claim that the meaning is constant, because it occurs only in one environment. The question which then arises is whether we can call *-ric* an affix. We might want to make the comparison with *bilberry*, where *bil* does not occur elsewhere, and with the difference between *ear* and *hear*, which might suggest an affix *h-*.

The difference appears to lie not in the examples themselves, but in words which seem to be parallel to the relevant examples. For example, parallel to *bishopric* we have words like those in (5), while parallel to words like *bilberry* we have words like those in (6).

- (5) baron-y, duch-y (where *duch* is an allomorph of *duke*), earl-dom, king-dom, princ-ipal-ity
- (6) black-berry, blue-berry, cloud-berry, goose-berry, snow-berry, straw-berry

In (5) the parallels are all with other words which do have recurrent affixes, in (6) the parallels (where they are clear) are with words that are compounds, while if we consider *hear* and try to make a parallel with words like *listen*, *see*, *sense*, *smell*, *taste*, *touch* we find no parallels at all. Therefore we decide that *bishopric* contains an affix, that *bilberry* does not contain an affix, and neither (for a different reason) does *hear*.

8.1.5 The Question of Affixoids

In the cases listed above it has been assumed that it is always possible (perhaps with a little bit of good will) to distinguish between a word and an affix. In some cases, however, this may not be true. An affixoid (sometimes called a “semi-affix”) is something that has an in-between status, or whose status is not readily determinable. Different definitions of affixoid can be found in the literature, some more restrictive than others (see for instance Marchand 1969, Fleischer 2000, Booij 2007, and Fleischer and Barz 2007: 27 for discussion). Here we take a relatively inclusive approach to affixoids.

An example can be found in Russian, where we find some elements borrowed from English whose status is extremely obscure. Their origin is in English words, but of course that does not necessarily imply that they are words in Russian. Typically these elements

occur in words which look like English compounds; but one of the elements (perhaps both) cannot occur in isolation in Russian, or if it can, it does not inflect, so that it is difficult to attribute it to any word class. At the same time, the semantics of this type is rather too specific to be typical of an affix. Some examples are presented in (7).

(7)

Russian element	Example	Gloss	Found in isolation?	Inflects in Russian?
art	art-magazin	‘art shop’	occasionally	not in isolation
audio	audiozapis’	‘audio recording’	no	no
kino	kino-kamera	‘film-camera’	only in informal styles	no
media	media-imperia	‘media-empire’	occasionally	no
meiker	tatu-meiker	‘tattoo-artist’	no	yes
pop	pop-versiya	‘pop-version’	no	no
super	superzvezda	‘superstar’	occasionally	no

(Liza Tarasova, personal communication)

The forms in (7) are in some ways parallel to neoclassical compounds in English: the elements are foreign and their manner of combination is also foreign, but they may be added to native elements. Borrowing is not the only way in which things classified as affixoids may differ from affixes or stems, though: typically semantic demotivation is used as a test. For example, the Dutch element *hoofd*, literally ‘head’ is used in a word before a noun to mean ‘chief, main’ (compare English *head office*) and might be analysed as an affixoid as a result. An analysis of an element as an affixoid suggests that it is not like typical affixes but not like typical bases, either.

8.1.6 Processes as Affixes

If we view all of morphology as being a matter of bases and affixes, we soon run into trouble, and English provides some good examples of such troubles. The difficulty is that in some cases it seems to be easier to view the difference between two morphologically related words as being determined by a process than by the addition of an affix.

Consider first the examples in (8) where the difference between a noun and its corresponding verb is a difference in the stress.

(8)

Noun	Verb
álloy	allóy
cónduct	condúct
cómpress	compréss
décrease	decréase
díscout	discóunt

éxploit	explóit
frágment	fragmént
ínsult	insúlt

There are over a hundred pairs like those in (8) in English. From time to time it has been suggested that there is an affix whose form is a particular stress pattern, but this seems to stretch the notion of an affix rather too much. It seems easier to say that there is process of assigning stress in the two cases, and that this process gives rise to the differences we see in (8). It also explains why some of the vowels that are not stressed are reduced to [ə], while others are not: the process affects vowels in different ways. (Note that the discussion above has not been formulated as deriving the noun from the verb or vice versa, which is what we would expect with an affix: you add an affix to a base. Most linguists see the difference in the forms in (8) as arising from the fact that nouns are stressed differently from verbs, according to different rules. This view makes it seem even less likely that an affixal analysis of the difference is justified.)

Next consider the items in (9). Here the difference lies not in the stress, but in the final consonants.

(9) <i>Noun</i>	<i>Verb</i>
belief /bili:f/	believe /bili:v/
house /haus/	house /haʊz/
mouth /maʊθ/	mouth /maʊð/
sheath /ʃi:θ/	sheathe /ʃi:ð/
shelf /ʃelf/	shelve /ʃelv/

In (9) the verb always has as its final phoneme the voiced equivalent of the voiceless phoneme that ends the noun. The rule is not a completely general one, but there is a larger class of affected words than is illustrated in (9). If we want to deal with this as a matter of base and affix, we have to say that the final consonants of both the noun and the verb are affixes, and the rest is the base. But this does not make sense of other affixation processes: *believable*, *mouthful*, and *mouthless*, and *shelving* are derived from the noun or the verb, and we can tell which, but they are not derived from a base missing a final consonant. A better analysis seems to be to say that there is a process which changes the final consonant as we move from noun to verb or vice versa. The same type of analysis seems to work with the examples in (10).

(10) <i>Noun</i>	<i>Verb</i>
song	sing
seat	sit

Again, it seems misleading to say that the base is *s_ng* rather than that the base is *sing*. This kind of process goes by many names: Ablaut, apophony, vowel alternation. But it is a process and not affixation in the middle of a base, because there is no parallel for that analysis and because it is contrary to the structure of bases in English, although there are

occasional mentions of these constructions as being infixes or infix-like in the literature (e.g. Kastovsky 2006c: 161).

The processes illustrated in (8), (9), and (10) cannot be affixes, because they do not fit the definition; they do have the same effect as affixes. In the past there has been some attempt made to fit examples like these into an affixal view of morphology, to make all of morphology have to do with bases and affixes (Nida 1948). Such analyses are now not favored because they stretch the definition of affix too far. Affixes might be extremely frequent, but they are not all there is in morphology.

8.2 THE POSITIONING OF AFFIXES

We have already seen that affixes can be attached to bases in different ways, although attention has not been drawn specifically to this point. In *king-dom* the affix comes at the end of the word, in *pre-paid* it comes at the beginning. We will see that there are other possibilities as well. More than this relative ordering of base and affix, we also need to draw a distinction between continuous and discontinuous bases and affixes. Continuous elements are not interrupted by other elements; discontinuous elements are.

8.2.1 Continuous Affixes

Continuous affixes can come before their bases, in which case they are called “prefixes,” after their bases, in which case they are called “suffixes,” or in the middle of their bases (so the base becomes discontinuous), in which case they are called “infixes.” An example of a suffix in Margi is given in (11), a prefix in Japanese in (12), and an infix in Khmer in (13).

(11) Margi

ḡàḡàl	‘hard’	ḡàḡàlkùr	‘hardness’
dènámá	‘strong’	dènámakùr	‘power, strength’
dúmi	‘bad’	dúmikur	‘evil’
dzáu	‘difficult’	dzáukùr	‘difficulty’

(Hoffmann 1963)

(12) Japanese

hizuke	‘date’	muhizuke	‘undated’
imi	‘meaning’	muimi	‘meaningless’
todoke	‘notification’	mutodoke	‘without leave’
kyooiku	‘education’	mukyooiku	‘uneducated’
zyooken	‘condition’	muzyooken	‘unconditional’

(Shibatani 1990, Martin 2004)

- (13) Khmer
- | | | | |
|-------|------------|---------|-----------|
| khos | 'be wrong' | kəmhos | 'a wrong' |
| khɤŋ | 'be angry' | kəmhɤŋ | 'anger' |
| sɔŋy | 'to speak' | səmdɔŋy | 'speech' |
| tba:ŋ | 'to weave' | təmba:ŋ | 'loom' |

(Jacob 1968)

Note that while the glosses given for the complex words in (11)–(13) are not consistent, since they depend on the vicissitudes of the English language, the semantic relationships are consistent: an abstract noun is created from an adjective in (11), a negative adjective is created from a noun in (12) which we might gloss as 'not having ~', and a noun is created from a verb in (13).

There are various points about (11)–(13) which it is worth making explicit. In each of these cases, the affix is made up of a whole syllable; this is not necessarily the case, as is illustrated by the difference between *warm* and *warmth* in English. The examples in (11)–(13) have been chosen to illustrate instances where the affix has a constant form, but affixes may take on several forms. Typically, the variation of shape between allomorphs is to allow ease of pronunciation over the base–affix boundary, and is determined by the phonology (the allomorphy is then said to be "phonologically conditioned"). There are also instances where the allomorphy is in principle unpredictable, at least from the shape of the surrounding elements, in which case it is lexically conditioned (that is, the allomorph is predictable only in terms of the individual base or the output lexeme). Examples of these phenomena from English are given in (14) and (15).

- (14) English: affixal allomorphy, phonologically conditioned

bed	embed	danger	endanger
body	embody	dear	endear
panel	empanel	list	enlist
power	empower	title	entitle

- (15) English: affixal allomorphy, lexically conditioned

adapt	adaptation
adopt	adoption
complete	completion
compete	competition

In the same way, the base may be modified under affixation so that there is base allomorphy. Some of this has already been illustrated in (15), but more examples are given in (16), where some of the variation in the base appears to be lexically conditioned, and other parts appear to be phonologically conditioned (albeit with phonological rules that apply only to suffixes of this kind).

(16)	compel	compulsion
	confuse	confusion
	define	definition
	dispose	disposition
	emerge	emersion
	induce	induction
	supervise	supervision
	tense	tension

Finally, note that with infixes it is not sufficient to know that there is an infix, it is also necessary to know where in the base the infix is to be inserted. In order for us to say we have an infix, it must interrupt some morph (usually the root, but sometimes another affix): an affix which occurs between two prefixes is itself a prefix.

There is a final kind of affix which has to be considered here, the interfix. As its name suggests, the interfix is placed between two bases. It is alternatively termed a linking element (Bauer 2003: 29–30). Consider the examples in (17) with German compounds.

(17)	Frau	'woman'	Zeitschrift	'magazine'	Frau-en-zeitschrift	'women's magazine'
	Kind	'child'	Spiel	'game'	Kind-er-spiel	'child's play'
	Liebe	'love'	Lied	'song'	Liebe-s-lied	'love song'

By saying that something is an interfix, we make certain assumptions about the analysis, namely that the affix belongs properly to neither side. It is always possible with an interfix to have an alternative analysis where the affix is a suffix to the left-hand element or a prefix to the right-hand element, and the suffixal analysis is frequently used in the German examples above (and corresponds to the historical development). Some authors see interfixes like those in (17) as empty morphs, that is meaningless affixes, on the grounds that there are, in German, parallel compounds such as *Zeit-geist* "time-spirit; defining mood of a period" with no linking element (Bauer 2003). It would also be possible to say that interfixes have a function which (in words like *Zeitgeist*) is not overtly expressed. Having no function or meaning would make them very atypical affixes (see the definition in Section 8.1.1 above).

8.2.2 Discontinuous Affixes

Other affixes are themselves discontinuous. They come in two or more parts, attached to the base in different positions. The parts of the affix can occur before the base, after the base, or in the base. Such affixes have been called "synaffixes" (Bauer 1988a), but are usually termed instances of "parasyntesis."

The type of parasynthesis that seems to gain most attention is the so-called “circumfix.” A circumfix is made up of two parts, one of which occurs before the base and the other of which occurs after the base. In convincing instances of parasynthesis, the two parts must make up a single affix, which is usually taken to imply that if we have a word of the form *X-base-Y* where *X...Y* is the circumfix, there is no semantically related form *X-Base* and no semantically related form *Base-Y*. A more restrictive requirement would be that there must never be any words of form *X-Base* or *Base-Y* which fulfill the same function as *X-Base-Y*.

Consider the case of German past participles. From the verb *mach-en* ‘to make’ we get the past participle *ge-mach-t*. Superficially, this looks like a circumfix, and is often analysed as such. But we also find verbs such as *ver-such-en* ‘to attempt’ whose past participle is *ver-such-t*. This could be taken as evidence that the *ge-* formative is not a necessary part of the past participle marker, and so more loosely related to the base than the final *-t*. Alternatively, it could be argued that the prefix *ver-* and the *ge-* are in a paradigmatic relationship, and the presence of one excludes the other. Even this seems to imply that *ge-* is a prefix in its own right, without the *-t*. The lesson to be learned from this example is that the existence of circumfixes is something which needs to be specifically argued for, and cannot be assumed without such an argument. Perhaps for this reason, really clear examples are rare.

There are also theoretical reasons for not accepting an analysis which depends on parasynthesis. Some scholars like to see morphological structure as being subject to the same kinds of rules that build up syntactic structure. Typically, this involves the ability to put a morphological structure into a tree structure. Where prefixes and suffixes are concerned, this creates no problem of principle. But it is also a requirement of some syntactic theories that trees should have a maximum of two branches. While a form like *mach-en* can be accommodated in such a tree structure, a form like *ge-mach-t* cannot be unless it is assumed that the *ge-* formative is added to the tree in a separate operation from the addition of the *-t*. So for such a binary analysis to hold of morphological structure, it has to be the case that circumfixes can be split into pairs of affixes, and that the order of affixation can be determined. Similar problems are associated with infixes. So at some point it may be necessary to determine which theoretical point has more weight: the requirement that trees are always binary or the apparent unity of a circumfix.

For an example of a circumfix, consider the data in (18).

(18) Cavineña

jutu	‘to dress.o.’	e-jutu-ki	‘cloth’
sama	‘to cure’	e-sama-ki	‘medicine’
taru	‘to stir’	e-taru-ki	‘paddle’
teri	‘to close s.th.’	e-teri-ki	‘door’
wijitu	‘to block s.th.’	e-wijitu-ki	‘stopper, cork’

(Guillaume 2008: 435–6)

Kambera			
ngùru	‘murmur’	ka-ngùru-k	‘to make a murmuring sound’
ndùru	‘thunder’	ka-ndùru-k	‘to make the sound of thunder’
mbàti	‘drip’	ka-mbàti-k	‘to make a dripping sound’
hètu	‘sniff’	ka-hètu-k	‘to make a sniffing sound’

(Klamer 1998: 245)

There are prefixes of form *e-* in Cavineña but their meanings do not seem related. So *e-* can mean ‘I’ or ‘noun,’ but the noun prefix is used mainly with parts of wholes and is deleted in the presence of other affixes. The suffix *-ki* by itself means ‘typical,’ or there is a suffix of form *-ki* which derives adjectives meaning ‘having ~.’ None of these seems relevant to the parts of the instrumental circumfix. In Kambera there are other instances of a prefix *ka-*, but no suffix *-k*, and the prefix *ka-* cannot attach to ideophones (as it does in (18)) when it occurs by itself.

Parasynthetic affixes are not necessarily placed round the base, they may interact with it in more complex ways. Some isolated examples are provided in (19). While all of these look as though they could be parasynthetic, it may be that the parts of the affix may recur independently of each other.

(19)

Jamul Tiipay	Miller 2001: 96	nyilly	‘be black’	taanyillya	‘blacken’
		kwelsaw	‘be clean’	kwellaasawa	‘to clean’
Tagalog	Miller 2001: 102	newill	‘forbid’ (sg)	anchuuwiill	‘forbid’ (pl)
	Blake 1925: 89	ibig	‘want’	kaibig’ibig	‘loveable’
		tákot	‘fear’	katakottákot	‘fearful’
	Blake 1925: 87	tábas	‘cut’	tinabásan	‘scraps of cloth’
Tzutujil	Dayley 1985: 176	káyas	‘pare’	kinayásan	‘parings’
		tik-	‘sow’	tijkoʔm	‘sowing’
		loq’	‘buy’	lojqoʔm	‘item bought’

Root-and-pattern morphology, dealt with in detail in Chapter 12, is sometimes dealt with as a complex pattern of infixation, in which case the term “transfixation” is used (Bauer 2003: 30).

8.2.3 The Psychology of Affix Position

There is some evidence that the most important part of a word for the hearer is the beginning: it allows the hearer to start processing the word and find the content in the

most efficient way (Cutler et al. 1985, Hall 1992). If the beginning of a word is a prefix, the processing may be slightly delayed, while a suffix does not delay the processing of the content of the word. This has been suggested as the reason for the fact that languages in general seem to prefer suffixes to prefixes: there are more languages that have suffixes but not prefixes than there are languages which have prefixes but not suffixes; in any given language which has suffixes and some other kind of affix, suffixes will be more common. This has been called the “suffixing preference” (Cutler et al. 1985).

8.2.4 The Typology of Affix Position

Many linguists suggest that the affix which determines the word-class of the word in which it appears (as does *-ness* in *sweet-ness*, for example, since it turns an adjective into a noun) is the “head” of the word (this notion will be discussed further in Section 8.4). That being the case, it might be expected that just as there are correlations between the ways in which heads of different constructions behave in syntax, so there would be correlations between the behavior of syntactic heads and morphological heads like *-ness*. Some figures are provided by Hawkins and Gilligan (1988: 228), who provide the data in (20).

(20) Syntactic form	Exclusively prefixing languages	Exclusively suffixing languages	Languages which have both prefixes and suffixes
VO	10	17	73
OV	0	62	38
Prepositions	7	21	72
Postpositions	1	65	34

That is, exclusively prefixing languages (ones with the morphological “head” on the left) tend to be those languages in which the verb precedes its direct object (syntactic head on the left) or in which there are prepositions (syntactic head on the left), whereas exclusively suffixing languages tend to have the syntactic head as well as the morphological “head” on the right.

The fact that there are more exclusively suffixing languages than exclusively prefixing languages might be explained by the processing difficulties outlined in Section 8.2.3. There is also another possibility. Defining the word is notoriously difficult, and one of the places where it is particularly difficult is where there are prefixes. Speakers of languages with suffixes tend to agree that the suffixes are part of the preceding word; speakers with prefixes may treat these prefixes as separate words. A nice example from two relatively closely related Bantu languages is provided in (22). The tense and agreement

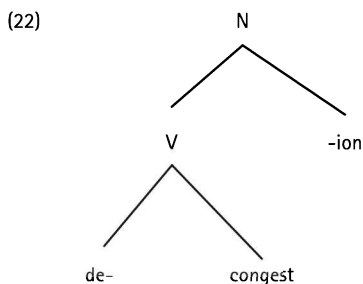
are marked as prefixes in Swahili, but as separate words in Kalanga. To the extent that the Kalanga type of analysis is widespread, it may be that prefixing languages are under-represented in the data.

- (21) a. Swahili
 Romeo a-na-m-penda Juliet
 Romeo he-PRES-her-love Juliet
 ‘Romeo loves Juliet’
 b. Kalanga
 Romeo u no da Juliet
 Romeo he PRES love Juliet
 ‘Romeo loves Juliet’
 (Fromkin 2000: 291)

8.2.5 Affix Ordering

When several affixes are all attached to the same base, the question arises as to how they are ordered in relation to each other. Clearly, if some of the affixes are prefixes and some are suffixes, as in *de-congest-ion*, the question may not arise, but when several prefixes or several suffixes are involved, there is a potential problem. So in *un-en-joy-able* and in *person-al-iz-ation* there are, in principle, questions to be answered, and several solutions are suggested in the literature.

Before we look at some of the proposals, two points need to be made. The first is that although this looks like a question of linear order, it may be reinterpreted as a question of hierarchical structure. It was mentioned earlier (Section 8.2.2) that bases and affixes may be viewed as being in a tree structure, so that *decongestion* gets the structure in (22). That being the case, linear order of affixes is related to tree structure.



The second point that has to be made is that ordering of affixes in derivational morphology is rarely contrastive. Given a word like *personalization*, it will rarely, in any language, be the case that there is another possible word *person-iz-al-ation* which will mean something different. At the same time, orderings are usually fixed, so that if *personalization* exists, *personizationalion* probably does not.

The first option is that affixes are selected by an adjacent affix, so that *-al* allows a following *-ize*, but not a following *-ation*. This suggests that not only every affix but also every base is idiosyncratic in what it allows in the adjacent position (Giegerich 1999). While there is some evidence to support such a notion (consider *bishop*, which is the only base to allow a following *-ric*, for instance or *hate* which is the only base to allow a following *-red*), in most instances this approach seems to specify more than is necessary.

A second approach is based on so-called stratal approaches or level ordering. In this approach, affixes are divided up into a number of sets (for English derivation, typically two, although larger numbers have been argued for from time to time—see the discussion in Giegerich 1999: 3), and those sets determine the relative order of affixes (see, for instance, Siegel 1979, Selkirk 1982). In English, affixes are divided into Class I affixes such as *a-*, *en-*, *in-*, *-able*, *-al*]_{Adj}, *-(a)tion*, *-ee*, *-ese*, *-ic*, and others and Class II affixes such as *un-*, *-al*]_{Noun}, *-er*, *-hood*, *-less*, *-ment*, *-ness*, *-y*]_{Adj} and others. Where strings of affixes occur, Class I affixes occur closer to the base than Class II affixes. So *un-* has to precede *en-* in *unenjoyable* because *un-* is Class II and *en-* is Class I; and in *sizeable-ness*, Class I *-able* has to precede Class II *-ness*. There are at least two problems with this approach. The first is that there are too many exceptions, like *contain-er-ize*, which break the rules for no obvious reason. The second is that there are many words with sequences of Class I affixes or sequences of Class II affixes which this approach does not provide any help with.

A third way of trying to provide generalizations in this area is to look at the affixes, and let them select what they will attach to (this is like the first approach, but the other way round: now affixes are looking to their bases rather than outwards to the next affix; Fabb 1988). Much of what was captured by the stratal approach can be captured here in terms of a split between native and learned vocabulary. In general terms, Class II affixes from the stratal approach are native, and Class I affixes are learned. But bases can also be native or learned. And learned affixes attach to learned bases in most instances, while native affixes attach to either native or to learned bases (or bases containing other affixes).

Somewhere in all of this, it is helpful to see particular affixes as attaching to bases of particular word-classes. So the suffix *-ation* in English attaches to verbs to produce nouns, and *-al* attaches to nouns to produce adjectives. The sequence of suffixes in a word like *operat-ion-al-ize* is partly determined by the sequence of word-classes (verb-noun-adjective-verb) that the affixes make the root undergo (Fabb 1988).

The order of affixes may also be determined by psycholinguistic factors such as parseability (Hay 2003). Other things being equal, affixes become more parseable the more of them that are added. Parseability includes semantic transparency (the meaning of the affix must be clear to the speaker and hearer) and phonological transparency (the affix must have an easily predictable form and must cause as little allomorphy in the base as possible; in the ideal situation, the phonotactics will also make it clear that there is an affix present). Consider the word *confid-ent-ial-ity*. There is allomorphy from *confide* so the *-ent* is not phonologically transparent; the meaning difference between *confident* and *confidential* is not predictable, so the addition of *-ial* is not semantically transparent;

but adding *-ity* to the end is both phonologically and semantically transparent. The outermost affix is the most transparent one.

Finally, there is the matter of semantic scope. If we go back to *person-al-ize-ation*, it is the act of personalizing; to personalize is the act of making personal; and personal means belonging to a person. So every affix adapts the meaning of the entire word to which it is added (Rice 2000). (There are occasional exceptions, but this general pattern is mostly true.) Each affix has scope over the entire base, including previously added affixes. This is true even where the affixes are not all on the same side of the base, as shown in (22). Thus the meaning determines the order of affixes.

Rather than any one of the factors that have been considered here being the overall determining factor in affix ordering, it seems likely that they all interact in some complex way to create sequences of affixes. Accordingly, affix ordering is an area in which considerable research is being carried out.

There is another alternative which has been left out of the above discussion because it seems to deal with a completely different way of viewing (and potentially processing) affix ordering. Some languages are said to have “templatic” morphology, where the word form fits into a template which specifies the set of possible affixal positions and the affixes which can be used in each position. An illustration of this phenomenon from Ket is provided in Chapter 15. It is sometimes the case with templatic morphology that the same form will have a different meaning when it occurs in a different position, and that individual positions are not realized in particular word-forms.

8.3 AN ASIDE ON REDUPLICATION

Reduplication (see Chapter 11 for fuller discussion) is sometimes viewed as a type of affixation. Bourchier (2008) notes that with reduplication, prefixation, and infixation are more common than they are with fixed-form affixation, and this might be an argument in favor of the position that fixed-form affixation and reduplication are not variants of the same morphological process type.

8.4 HEADEDNESS

The notion of headedness comes from syntax. In syntax the head of a construction is the obligatory element in a construction which defines the type of construction to which it belongs and which is the semantic defining element of the construction, which other elements modify. So in a noun phrase like *interesting books about linguistics*, *books* is the head because all the other elements can be deleted but *books* is obligatory, *books* defines the phrase as being a noun phrase and *books* is a superordinate term for *interesting books about linguistics*. Similarly, if we move from syntax to word-formation, in a compound

such as *whiteboard*, *white* can be deleted but not *board*, *board* defines the compound as being a noun and *board* is a superordinate term for *whiteboard*; the conclusion is that *board* is the head in *whiteboard*. The question is whether and how this notion applies to derivation.

Where English is concerned, we need to consider three distinct types: words like *white-n*, words like *un-fair*, and words like *un-sex*. If we consider each of these three words against the three criteria that were listed above, we find the results set out in (23). In (23), *a* means “affix” and *b* means “base.”

(23) The head of a word: contrasting results from different criteria

	Obligatory	Word-class definition	Semantic center
whiten	b	a	?
unfair	b	b	?
unsex	?	a	?

This table requires some explanation. The obligatoriness criterion does not work well with derivatives. The base is obligatory in an instance like *whiten* for two reasons: we cannot have an affix without a base, and some bases like *yellow* (see (4) above) have no affixes to produce a verb. In a case like *unfair*, *fair* is obligatory in precisely the same sense that *books* is obligatory in *interesting books about linguistics*: all the other elements can be omitted and what is left will still be grammatical. In *unsex*, only one of these arguments holds: we cannot have an affix without a base.

In *whiten* it is the *-en* affix which defines the word as a whole as a verb. In *unfair*, the whole word is an adjective because *fair* is an adjective. It has been argued (e.g. Williams 1981b) that these instances generalize as the “right-hand head rule,” that the element on the right-hand branch of any morphological tree is the head. An example like *unsex* seems to provide counter-evidence to this generalization. *Sex* is a noun, and *unsex* is a verb because of the prefix, which is on the left-hand branch of the tree. It might be possible to argue for a zero-affix on the right to head a construction like *unsex-Ø* (as argued by Nagano 2011), but *un-* is not attached to nouns with the relevant meaning and there is no relevant verb *sex-Ø* “to give someone the attributes of a sex.” The arguments against zero-morphs (see Section 8.1.3) are also relevant here.

Where the semantic argument from hyponymy is concerned, it is unclear what we are to conclude. *Whiten* is not a hyponym of *white*, but could be taken to be a hyponym of an element which means “causative verb.” *Unfair* is not a hyponym of *fair* (just the opposite, it is an antonym), but it is also hard to see it as a hyponym of *un-*. Similarly with *unsex*: the verb *unsex* cannot be a hyponym of the noun *sex*, but not everyone is happy with the notion that it can be seen as a hyponym of *un-*, which we would have to gloss as ‘to remove features connected with ~ from.’

The only reasonable conclusion that can be reached from such examples is that there is nothing in derivation which is just like a syntactic head. However, scholars who have

made a great deal from models of morphology that include headedness (e.g. Lieber 1992) use only one of the three criteria listed in (23), namely the word-class criterion. At this point, headedness is defined in a particular model, and the results produced by that model are the justification for the definition. It just seems unfortunate that the morphological “head” defined in this very narrow way continues to be equated with the syntactic head where head has a different definition.

8.5 THE MEANINGS OF AFFIXATION

Bybee (1985) notes that the meanings of affixes have to be sufficiently general to apply to a wide range of bases, otherwise there is no reason to have an affix. Another way of looking at this is that, in general terms, affixes provide grammatical or functional rather than lexical meaning. One of the reasons that neoclassical word formation is usually considered to be compounding is because each element has lexical value: *morph-o-logy* means ‘the study of form’ where both elements are glossed in English with full lexemes. Even with this starting point, however, it can be difficult to decide just what kind of meaning an affix might be expected to carry.

A prefix like *un-* carries a meaning of “negative,” and that is typically seen as grammatical meaning. It is certainly general enough to apply to a large number of bases. Prefixes like *pre-*, *post-*, *trans-* have meanings, respectively, ‘before,’ ‘after,’ and ‘through/across,’ where the glosses are prepositions, and thus the meaning can again be seen as being grammatical. However, the suffix *-er* has a host of meanings, among which we might single out ‘agent,’ ‘patient,’ ‘instrument,’ and ‘location’ as in *killer*, *boiler* (‘boiling fowl’), *cooker* (‘cooking stove’), and *diner* (‘place to eat’), respectively. It is less clear that these are grammatical, although they can also be marked in syntax by the prepositions *by* and *in*, for instance. The meaning ‘female’ associated with the suffix *-ess* in *princess* might appear grammatical in languages which have grammatical gender, but in English, this seems rather more lexical. And the affixes *step-* in *step-mother* and *-iana* in *Victoriana* seem to be extremely lexical and surprisingly non-general.

We must also recall that the meaning of an affix should be distinguished from the meaning that an affix carries in a particular word. There are languages which have affixes which only ever mean ‘agent,’ for instance, but given that *-er* in English can mean ‘agent,’ ‘patient,’ ‘instrument,’ or ‘location,’ it seems more reasonable to say that the meaning of *-er* is actually more abstract than any of these (Lieber 2004), but is made up of what these have in common: perhaps they do no more than denote an argument of the verb in the base, for example, or, if we include constructions illustrated by *tenner* and *cold mooner* as illustrating the same affix, even more abstract than that. Alternatively, it would be necessary to postulate a large number of homophonous affixes, an approach which seems to garner little support.

What we see, then, is a cline from the more canonical grammatical and general affix-like meanings to rather more lexical meanings. It should also be borne in mind

that what counts as “general” may be, at least to some extent, culturally determined. There are Australian languages which have a suffix meaning ‘the person who died at the place mentioned in the base,’ a culturally important category (e.g. Mangarayi, see Merlan 1982), but not one that is important enough in Anglo-Celtic culture to have an affix in English. How far this notion of cultural importance can be pushed is unclear, though. There are languages with very few derivational affixes, but it would be an error to say that such a language has few culturally important categories. Consider the notion of “collection of” in English, for example. Although we have a host of lexemes to denote the notion (*bevy of quail*, *kindle of kittens*, *parliament of owls*, *pride of lions*, *school of dolphins*, *swarm of bees*, and so on), which suggests that the category is important for English speakers, there are relatively few words for such collections which are marked by affixes (*rookery*, *swannery*).

8.6 CONCLUSION

Affixation is a way of making new words from old by adding phonological material to the original. It is concatenative. There are many other aspects to morphology beyond affixation, but affixation is probably the default type of structure-building in morphology. The meanings of affixes in derivation vary from the extremely grammatical to the very lexical; the ordering of affixes is constrained in many ways which are not yet fully understood, although many basic principles are becoming clear; just what counts as an affix and what does not may be controversial.

The discussion here has, because of the nature of the book, been constrained to instances where derivation is at stake. Affixation is also used in inflectional morphology. The main difference is that with inflectional affixation care has to be taken to distinguish it from cliticization, while with derivational morphology, care has to be taken to distinguish it from compounding. Inflection and derivation also have to be distinguished in those theories that see an important difference between them (see Chapters 2 and 3).

CHAPTER 9

INFIXATION

JULIETTE BLEVINS

9.1 DEFINITIONS: INFIXATION AND DERIVATION

A chapter on infixation in a volume on derivation should narrow down the type of word-formation processes under study. However, the terms “infixation” and “derivation” have a wide range of meanings in the linguistics literature. For this reason, it will be useful to clarify, from the start, how these terms are used here.

9.1.1 Infixation

When a word is readily analyzed into subparts with clear form–meaning correspondences, we say that it is morphologically complex. Infixation is morphological complexity of a very specific type, as defined in (1). Under infixation, a base, be it a root, stem, or word, is phonologically discontinuous due to the presence of an infix *inside* it.

(1) Infixation as affixation: a definition

Under infixation a bound morpheme whose phonological form consists minimally of a single segment, is preceded and followed in at least some word-types by non-null segmental strings which, together, constitute a relevant form–meaning correspondence of their own, despite their non-sequential phonological realization.

There are four factors that distinguish this definition from others. Each will be briefly reviewed and justified here. Following the Leipzig Glossing Rules, infixes and infixed material are enclosed in angled brackets < > in glosses and citations.

First, the definition specifies that an infix is a bound morpheme, falling within the general definition of an affix. While free forms, including multi-morphemic strings, can also split roots, stems, and words into discontinuous parts, this process is referred to as *tnesis*, and is distinct from infixation as discussed here. A canonical instance of true infixation is shown in (2) from Hoava [hoa], an Oceanic language of the Solomon Islands (Davis 2003).¹ The infix in question is <in>, a productive nominalizer that can be used with “virtually any active or stative verb to create a noun,” and which only occurs as a bound form (Davis 2003: 39). The infix is consistently placed before the first vowel of the base.

(2) Infixation in Hoava: <in> nominalizer, a bound morpheme

Locus: precedes first vowel of base

<i>root</i>	<i>stem</i>	<i>infix</i>	<i>glosses</i>
a. to		t<in>o	‘alive/life’
b. hiva		h<in>iva	‘want/wishes’
c. bobe	va-bobe	v<in>abobe	‘full/fill/filled object’
d. poni	ta-poni	t<in>aponi	‘give/be given/gift’
e. razae	vari-razae	v<in>arirazae	‘fight/fight each other/war’
f. asa		<in>asa	‘grate/pudding of grated cassava’
g. edo		<in>edo	‘happy/happiness’

Compare the inserted bound morpheme in (2) with instances of *tnesis*, as illustrated in (3) by examples of expletive “infixation” in English (McCarthy 1982) in standard orthography, with primary stress marked in base forms. Under *tnesis*, a word is split apart by another word (free form).

(3) *Tnesis* in English: expletives render words discontinuous

Locus: before a stressed trochaic foot

<i>Base</i>	<i>Tnesis</i>
a. abso'lutely	abso<damn>lutely
b. fan'tastic	fan<fuckin'>tastic
c. rhe'torical	rhe<goddamn>torical
d. ty'phoon	ty<bloody>phoon

While the split may be phonologically conditioned, and therefore share phonological properties with true infixes, it differs morphologically from true infixation in two important ways. First, the “infix” is not a bound morpheme (e.g. *damn*), and can be a morphologically complex word itself (*god-damn*, *blood-y*, *fuck-in*). Second, forms derived from *tnesis* are neither inflected forms of the same lexeme, nor members of the same word-family: they are, in a sense, the *same word*, interrupted, or splattered in

¹ ISO language codes are those of Lewis (2009).

speech. Tmesis will not be considered further, though such cases are included in many phonological studies of infixation, most often with reference to the specific locus of the inserted word (e.g. Yu 2007a, b).

A second notable feature of the definition in (1) is the lack of restriction on a morphological base. Though some definitions of “infix” limit bases to roots, there is good evidence that the base of infixation may be a stem or word. The Hoava examples in (2) illustrate that “root” is too limiting a domain for infixation. In (2c–e), the phonology and semantics are consistent with infixation into a derived stem or word. In (2c) the root is a stative verb, /bobe/ ‘full,’ prefixed with causative /va-/ to mean ‘fill’; the nominalization of this verb is a “filled object.” Likewise, in (2d), it is not the root /poni/ ‘give’ that is the base of infixation, but the derived passive stem /ta-poni/ ‘be given’ that is nominalized, resulting in the compositional semantics of “that which is given, gift.” In (2e), the root /razae/ takes a reciprocal prefix /vari-/ , and this prefixed form is nominalized, resulting in the interesting morphotactic of one bound morpheme inside another. While none of these examples allow a clear distinction between stem and word, there is little question that the base of infixation in Hoava is not limited to roots, and must at least include derived stems.

However, in other languages, there is evidence that a single infix may take either a stem or inflected word as its base. In Yurok [yur], an Algic language of northwestern California, an intensive infix <eg> regularly follows the first consonant or consonant cluster of the base, as shown in (4), where *hl* writes /ɬ/, *g* writes /ɣ/, *c* writes /tʃ/ and *y* writes /j/ (Robins 1958, Garrett 2001, Wood and Garrett 2001).

(4) Infixation in Yurok: <eg> intensive, a bound morpheme

Locus: follows first C or C-cluster of base

<i>root/stem</i>	<i>infixated stem</i>	<i>glosses</i>
a. laay-	l<eg>aay-	‘pass’/‘pass regularly’
b. kemol-	k<eg>emol-	‘steal’/‘be a thief’
c. cwin-	cw<eg>in-	‘talk’/‘act as go-between in marriage negotiations’
d. hlk-	hlk<eg>-	‘gather acorns’/‘gather acorns regularly’

We will consider the derivational vs. inflectional status of this morpheme shortly. For the moment, what is of interest is the fact that, for Yurok *o*-stem verbs, verbal inflectional suffixes are sensitive to the phonological composition of the base (Robins 1958: 34, Blevins 2005). Indicative inflections for subsyllabic verb stems lacking vowels, like /hlk-/ in (4d), have long vowels in all but 3sg forms: *hlko:k*’ 1sg, *hlko:m* 2sg, *hlko*: 1pl, *hlko:w* 2pl, *hlko:hl* 3pl, in contrast to other *o*-stems which show short vowels throughout the paradigm. Since this vowel length difference in the inflected verb depends on the phonological composition of the stem, it allows us to determine whether <eg> can ever take a word as its base. Indeed, this appears to be possible, as illustrated by the two inflected variants shown in (5). Example (5a) is the expected case, where the infixated base appears to be the stem for inflection; the infix supplies a stem vowel, so the stem is monosyllabic,

and the final inflectional vowel is short. In contrast, (5b), which is also attested, is unexpected: here, the infixed 1sg form has the same inflection as the non-infixed subsyllabic stem /hlk-/, with a final long vowel. Since long vowel inflections appear only when the verb stem is vowel-less, the base for inflection must be /hlk-/, meaning the base for infixation is the inflected word *hlko:k'*.

(5) Yurok: <eg> intensive variants: base = stem *or* inflected word

<i>base</i>	<i>infixed base</i>	<i>1sg indicative inflection</i>
a. hlk-	hlk<eg>-	hlkegok' (-ok' w/ bases of at least one syllable)
b. hlko:k'	hlk<eg>o:k'	hlkego:k' (-o:k' w/subsyllabic bases only)

A third component of the definition in (1) which needs to be specified is the phonological requirement that infixes be minimally mono-segmental. A single vowel or consonant is not only a feature matrix, but one anchored in the timing tier, however this is represented in one's phonological model. Floating features, floating feature complexes, and unanchored segmental melodies do not satisfy the definition of full-blown segments, as they lack timing units, and so cannot constitute proper infixes in the sense defined here.²

Finally, though it is a minor point, note that the definition in (1) states that the discontinuity of the base, being broken into parts which themselves do not constitute sound-meaning correspondences, need only hold for "some word-types." For Hoava, word-types (2f,g) are vowel-initial, and therefore do not involve base-discontinuity, but this does not rule out <in> affixation as true infixation because C-initial words like (2a-d) do show discontinuous bases. This wording also allows for chance word-types where an infix appears between two distinct morphemes as in Yurok (5a,b); the critical observation is that in some Yurok word types (4a-c) the same affix gives rise to discontinuous meaningful units.

Infixes, like other bound morphemes, need not have a clear meaning associated with them, as discussed further in Section 9.5, and may qualify as "empty" morphs. They should not be confused with interfixes, also referred to as linking elements, connectives, linkers, or linking morphemes. Interfixes are meaningless morphemes that consistently occur between, not inside of, other morphemes (for examples, see Štekauer et al. 2012: 199–200).

9.1.2 Derivation

As this volume focuses on derivational morphology, we will limit discussion, for the most part, to instances where the infix in question is "derivational" as opposed to

² Floating features and feature-complexes, sometimes referred to as featural affixes, include labialization ([+round]) marking 3rd person masculine singular objects in Chaha and palatalization ([+high, -back]) marking 3rd person singular in Ithmus Mixe (Akinlabi 2011, Blevins 2012). Unanchored segmental elements include the well studied consonantal roots of Semitic languages (McCarthy 1981).

“inflectional” in the sense outlined in, for example, Stump (2005) and Chapters 1 and 2 of this volume.

Stump (2005) speaks to practical criteria for distinguishing derivation from inflection, and we use these throughout in a fairly conservative manner. If infixation imposes part-of-speech membership, then we treat it as derivational. If an operation is complete and semantically regular, it is usually inflectional, not derivational. If it is syntactically determined, it is also inflectional, not derivational. The one practical criterion we do *not* use is the structural one, which suggests that in general, marks of inflection are peripheral to marks of word formation; or in operational terms, derivational operations apply before inflectional ones. Because infixes, by their nature, require phonological locus-placement information about the relevant base as part of their lexical entries, they may have different structural properties from affixes that are aligned at the beginning or edge of a base.

As already shown for Yurok <eg> intensive illustrated in (4) and (5), there appear to be cases where an infix takes an inflected word as its base. If <eg> is derivational, this is a clear counter-example to the structural generalization, since, operationally, derivation follows inflection. But is Yurok <eg> an instance of derivation or inflection? As Stump (2005) illustrates, there are problems with the practical criteria at every step, and the Yurok intensive is no exception. While it is often used in the formation of deverbal nouns, its primary function is linked to repetition of an active verb or intensity of statives (Wood and Garrett 2001). Since there are semantic irregularities, specialized meanings (e.g. 4b, d) and a wealth of lexicalized deverbal nouns (e.g. *naʔaw*- ‘to catch surf fish,’ *nega* ‘surf-fish net’: *swehlk*- ‘to burst,’ *swegehl* ‘gunshot’), and since <eg> is not syntactically determined, derivation remains an option. In this case, like many others, however, it is a holistic view of Yurok grammar including its history, that may be more informative than practical criteria: as the inflectional system has undergone significant changes in line with clearly defined paradigms, <eg> has led a life of its own, suggesting, indeed, that it serves as a unique form of derivation in this language (Garrett 2001, Blevins 2005, forthcominga).

Though many infixes identified fail to satisfy all of the practical criteria for classification as inflection or derivation, I have tried, as far as possible, to eliminate these from discussion. Unless otherwise noted, in this chapter, an infix is classified as derivational when: (i) it does not express inflectional features of tense/aspect/mood, agreement, and/or case; and (ii) it may involve a change of grammatical category; and/or (iii) it is, for these and/or other reasons, classified by language specialists as a clear case of derivational morphology in a particular language.

9.2 PHONOLOGY AND MORPHOLOGY INTERACTIONS: A BRIEF SUMMARY

Detailed cross-linguistic studies of infixation include Ultan (1975), Moravcsik (2000), Yu (2007a, b), and most recently, the typological overview in Štekauer et al. (2012:

197–203). None of these studies focuses exclusively on infixation as defined in (1) above, nor is there a clear focus on derivation, as opposed to inflexion, but as many of the generalizations noted in these works cover derivational infixes as defined here, they are worth summarizing.

The definition of infix presented in (1) requires that an infix constitute at least one full-blown segment, be it a vowel or consonant. Infixes then, unlike morphemes generally, have a minimal size. The minimal form of an infix, then will be a single segment, and indeed infixes of this type, like Semai [sea] <r> causative, Quileute [qui] <j> diminutive, or Tausug [tsg] <i> nominalizer are common. As with other bound morphemes, infixes tend to be short, but there seems to be no clear phonotactic upper limit on size.

One phonological variable that infixes share with other bound morphemes is whether their segmental content is fully specified, partially specified, or unspecified. In the last two cases, reduplication is involved: language-specific mappings will determine the base and direction of melodic matching/copying for the unspecified reduplicative infix. There are many languages with reduplicative infixes that lack other infixes, a fact that is simply explained by their origins: historically, reduplicative infixes are reanalyses of reduplicative adfixes that have mutated over time (Yu 2007a: 165–70).

A unique feature of infixes, is that, by definition, there must be a specification in their lexical form of their precise infixation site with respect to the base. Where a prefix aligns with the beginning of a base, and a suffix with the end, more needs to be said for infixes. For example, in (2), the infix precedes the first vowel of the base, in (3) it precedes a stress-foot, and in (4) it follows the initial consonant or consonant cluster. An overarching generalization is that all infixes align themselves in some way to the beginning or end of the base,—what Yu (2007a, b) terms the “Edge-Bias Effect,” and that at each edge a set of phonological pivots define infixation sites. Within Yu’s model, there are two basic kinds of pivots: edge pivots and prosodic pivots. The edge pivots include the initial and final C, V, or syllable of the base; while the prominence pivots are the stressed vowel, syllable, or foot.

Within Optimality Theory, however, it has been argued that infixes and infixation are derivative notions: the only adfixes are prefixes and suffixes, aligned with the beginning and end of the base (Prince and Smolensky 1993; McCarthy and Prince 1993b). When an infix surfaces, it is due to phonological constraints dominating morphological specifications, essentially driving the adfix inside the base, to a position where phonological constraints are best satisfied. For example, the Hoava infix <in> in (2), would be treated as a simple prefix on the basis of forms like (2f,g), with apparent infixation in (2a–e) accounted for by phonological constraints like ONSET, which would give preference to well-formed *ti.no* ‘life’ where both syllables have onsets, over unattested **in.to*, with an initial onsetless syllable. Counter-evidence to this position from Leti [liti] is found in Blevins (1999), and further evidence for infixes as bound morphemes distinct from prefixes and suffixes is offered in Yu (2007a, b). The most persuasive argument for infixes as true infixes, are languages in which homophonous strings constitute distinct prefix/infix or suffix/infix pairs. Atayal [tay] <m> actor focus, and /m-/ reciprocal/reflexive prefix, may be just such a pair, as illustrated in (6), though it should be noted that both

the morphological status of /m-/ as marker of reciprocal/reflexive and the phonological status of <m> may be contested.³

(6) Atayal <m> vs. /m-/: a minimal infix/prefix pair (Egerod 1965: 266–7)

Root	Gloss	<m> Actor focus	/m-/ Reciprocal/ Reflexive
kaial	'talk'	kmaial	mkaial
qul	'snatch'	qmul	mqul
siuk	'give back'	smiuk	msiuk
sbil	'leave behind'	smbil	msbil

Another argument against infixes as phonologically optimally placed adfixes comes from languages where infixes trigger phonotactic repairs. One of the most remarkable cases of this kind is found in Arara [arr], a Cariban language of Brazil, where special forms of speech are used to talk to pet animals of different types (De Souza 2010). These special forms include regular prefixation and infixation, and depend on the type of pet one is talking to. Words uttered to pet squirrel monkeys are infixed with <pt>, as illustrated in (7), with this infix following the first vowel of the base. Arara maximal syllables are CVC, with V, CV, and VC all possible. While infixed forms like (7a,b) result in well-formed syllables, those in (7c–f) do not, and repair strategies are in evidence. In (7c,d) a copy vowel is inserted to break up the illicit word-final cluster, while in (7e,f) the medial cluster undergoes reduction from $C_1C_2C_3 > C_1C_2$.

(7) Infixation in Arara: <pt> in pet squirrel monkey talk

Locus: follows first vowel of base

base	infix	gloss
a. ae	a<pt>e	'wasp'
b. pou	po<pt>u	'small peccary'
c. nu	nu<pt>u (*nupt)	'small tumour'
d. wot	wo<pt>ot (*woptt)	'fish'
e. pitot	pi<pt>ot (*pipttot)	'a fruit'
f. abat	a<pt>at (*aptbat)	'manioc bread'

While one might argue that pet talk in Arara is a speech disguise, and therefore, not subject to grammatical constraints ranking phonological over morphological conditions, the fact remains that a word-formation process exists in this language where a specified phonological string is inserted *inside a morpheme*, with no phonological

³ One could analyze /m-.../ as underlying /p(ə)<m>.../ with (irregular) cluster reduction; and the infix itself could be analyzed as underlying <əmə>, not <m> (Rau 1992, Kaufman 2003).

motivation. Furthermore, this infixation results in poor phonotactics, which are remedied by repair strategies suggesting that words of the special language do indeed have high-ranking phonological constraints. Since similar instances of word-formation are found in common language (Blevins 1999, Yu 2007a, b), infixes must be posited as primary morpheme types, and infixation must exist as a derivational process distinct from adfixation.

In the rare cases where multiple infixes occur within a single word, derivational infixation appears to precede inflectional infixation and be closer to the stem. One example of this is found in Begak [dbj], a language of Sabah, where derivational <ər> reciprocal (8a,b), can co-occur with the inflectional completive aspect <ən> (8c) (Goudswaard 2005). Since both infixes take as pivot the first vowel of the base, it is clear that in *sənəratu*, inflectional <ən> is infixed to the derived reciprocal stem *s<ər>atu*. However, although this expected structural relationship holds of derivational and inflectional infixes, other structural relationships are found for derivational <ər> and inflectional prefixes: (8d) is again the expected case, where derivational <ər> derives a stem to which /gə-/ a dynamic transitive actor voice marker is prefixed; but in (8e), derivational <ər> appears unexpectedly inside the actor voice non-volative prefix /kə-/, suggesting that the base for reciprocal formation can be an inflected form.

(8) Begak double infixation: derivation inside inflection

Locus: <ər> reciprocal, follows first vowel of base

Locus: <ən> completive aspect, follows first vowel of base

	<i>base</i>	<i>reciprocal</i>	<i>inflected reciprocal</i>	<i>glosses</i>
a.	kanut	k<ər>anut	'pull'/'pull each other'	
b.	kati	k<ər>ati	'tease'/'tease each other'	
c.	satu	s<ər>atu	s<ən><ər>atu	'be one'/'be together'/'put together (completive aspect)'
d.	tabang	t<ər>abang	gə-t<ər>abang	'help'/'help each other'/'help each other.AV'
e.	kə-nnik	k<ər>ə-niik	(k<ər>ə-niik)	'AV.NV-ascended'/'ascended together'

Generalizations have also been made about the origins of infixes. Some derivational infixes, like deverbal Hoava <in> (2), are reflexes of infixes that have persisted across thousands of years (see Section 9.5), while others, like Yurok intensive <eg> appear to be relatively recent innovations (Garrett 2001). Where infixes are not directly inherited, a limited number of evolutionary pathways have been proposed for their evolution from adfixes, including entrapment, phonological metathesis, and reduplicative mutation. For further details, see Yu (2007a, b).

9.3 MEANINGS ASSOCIATED WITH DERIVATIONAL INFIXES⁴

Cross-linguistically, derivational morphemes are known to express a wide range of meanings, from the nuanced difference between English *green/greenish*, to category-changing functions like nominalization, and highly lexical meanings like Halkomelem /-wət/ ‘canoe.’⁵ The same is true for derivational infixes. Table 9.1 provides representative examples of non-reduplicative derivational infixes organized in terms of their semantic content. At the top of the table are infixes with negligible, bleached, or intangible meanings, followed by those with grammatical meanings, and finally, a few with lexical meanings.

Table 9.1 Semantics of derivational infixes: from intangible to lexical

	Language/Family/	Infix	Form/gloss
Intangible	Jeh/Austro-Asiatic	<an>	kriem ‘crossbow’ kadriem ‘crossbow’
	Thai/Tai-Kadai	<amn>	ʔuay ‘to bestow’ ʔamnuay ‘to bestow’ (elegant)
	Arara/Cariban	<pt>	ae ‘a wasp’ apte ‘a wasp’ (when talking to a pet squirrel monkey)
Grammatical	Alabama/Muskogean		hocca ‘shoot’ holicca ‘be shot’ (MEDIOPASS)
	Chamorro/Austronesian	<in>	hasso ‘think’ hinasso ‘thought’ (NOM)
	Semai/Austro-Asiatic	<r>	kdey ‘not to know’ krdey ‘cause not to know’ (CAUS)
	Pingding Mandarin/ Sino-Tibetan	<ɿ>	xua ‘flower’ xɿua ‘little flower’ (DIM)
Lexical	Klallam/Salishan	<aʔ>	čə’səʔ- ‘two’ ‘čáʔsaʔ ‘two people’
	Scientific English/ Indo-European	<pe>	lutidine
			lu<pe>tidine

⁴ Data sources and ISO codes for languages mentioned in this section are: Alabama [akz] (Martin and Munro 2005); Arara [ara] (De Souza 2010); Chamorro [cha] (Topping 1973); Halkomelem [hur] (Gerds 2003); Jeh [jeh] (Gradin 1976); Khmer [khm] (Huffman 1986); Klallam [clm] (Charles 2012); Mangarayi [mpc] (Merlan 1982); Pingding Mandarin [cmn] (Lin 1989); Semai [sea] (Kruspe 2004: 134); Shuswap [shs] (Kuipers 1974); Tetun [tdt] (Williams-van Klinken 1999); Thai [tha] (Huffman 1986).

⁵ For an overview of lexical affixes in Salish languages, see Czaykowski-Higgins and Kinkade (1998: 25–7).

The majority of derivational infixes are grammatical morphemes—most commonly nominalizers, verbalizers, causatives, diminutives, and intensives. When these morphemes become lexicalized, the semantic value of the infix in derived words moves towards the intangible end of the spectrum. The Jeh forms in Table 9.1 illustrate common lexicalization of a once-productive infix. The two other examples of infixes are very different in nature. As discussed in Section 9.7, Thai acquired infixes via contact with neighboring Mon-Khmer languages, but with semantic change. Where the infix <amn> functioned as a nominalizer in Khmer (used with bases beginning in single consonants), the same infix in Thai is used to create a stylistic variant of the base (Huffman 1986: 201–2). Another unique infixation, partly illustrated in (7), is that found in Arara pet animal talk (De Souza 2010). Recall that words uttered to pet squirrel monkeys are infixed with <pt>, as in (7) and Table 9.1, but in speech to a pet capuchin monkey, the partial reduplicative infix <gV> is used instead, and a range of prefixal forms are used for other pet animals, e.g. /pi-/ for words spoken to pet agoutis.

At the other end of the scale, there are very few reported derivational infixes with lexical meanings. This is not altogether surprising, since lexical affixes of the kind found in, for example, Salish and Eskimo-Aleut are rare cross-linguistically to begin with. As the majority of non-reduplicative infixes evolve from prefixes or suffixes via entrapment or metathesis (Yu 2007a, b), lexical infixes are expected in languages with lexical affixes from which they can originate. It is not surprising then that Klallam, a Salish language with a wealth of lexical suffixes, has at least one instance of a derivational infix, <aʔ> ‘person,’ with a highly lexical meaning.

The other example provided in Table 9.1 of a lexical infix comes from the scientific vocabulary of English, where a chemical, *picoline*, for example, is related to a derived form *pipecoline*, by infixation of <pei> meaning ‘completely hydrogenated.’ The source of this word-formation process seems to be analogy with the pair *pyridine*/*piperidine*, since *piperidine* is produced by the hydrogenation of *pyridine*, though, historically, *piperidine* is derived from Latin *piper* ‘pepper.’

One meaning-based generalization that does appear to hold of infixation is that reduplicative infixation is more likely to have the semantics associated with reduplication more generally, than with semantics of non-reduplicative morphemes. In other words, reduplicated infixes show higher frequency associations with plural, pluralational, intensive, repetitive, iterative, distributive, and augmentative/diminutive meanings than non-reduplicated infixes (cf. Rubino 2011). This association undoubtedly stems from the fact that reduplicated infixes result from historical reanalysis of earlier reduplicative prefixes and suffixes (Yu 2007a, b). Since most of these meanings are those associated with inflectional features of number and tense/aspect/mood, the majority of reduplicative infixes are non-derivational. When only derivational reduplicative infixation is considered, the database of languages is greatly reduced, but one still finds a range of meanings including: Chamorro <CV> adjectival intensifier; Tetun <aC> deadjectival nominalizer; Shuswap <C> diminutive; and Mangarayi <VC> intensive.

Since, overall, infixes are less common than prefixes and suffixes (Ulan 1975), it is not surprising that meanings associated with derivational infixes are a subset of those found for derivational prefixes and suffixes.

9.4 SOME QUESTIONABLE UNIVERSALS

Infixes are less common than prefixes and suffixes, but given their common occurrence in Austronesian (Blust 2009), a family with approximately 1,000 languages, and their attestation in at least 25 other phyla and isolates (Yu 2007a), they cannot be considered rare. Nevertheless, early proposals regarding their cross-linguistic distribution seem to assume that, along with circumfixes and templatic morphology, infixation is particularly marked.

One implicational universal along these lines that is often repeated is Greenberg's (1963b: 73) claim that the presence of "discontinuous affixes" (=infixes, circumfixes, or intercalated morphemes) in a language implies the presence of suffixes and/or prefixes; where affixation is at issue, there are no languages that employ infixation exclusively. In Greenberg's (1963b: 73) words, "If a language has discontinuous affixes, it always has either prefixing or suffixing or both," and a modern typological restatement "Variation is reined in by this implicational constraint: If there are infixes, there will also be adfixes (= suffixes and/or prefixes)" (Plank 2007: 58).

While there appear to be very few languages that have infixation, but lack other affixation types, such languages do exist.⁶ Pingding Mandarin appears to have only one affix, a diminutive/hypochoristic infix, </>, with pairs like *xua* 'flower', *x/ua* 'little flower' (Lin 1989, 2004, 2008). There are no other clear affixation processes in the language; all other word formation appears to involve compounding or cliticization. In this case, an earlier stage of the language is represented by Standard Mandarin, where the cognate morpheme is /-r/, a suffix with similar function, and phonological metathesis is responsible for the evolution of the infix (Yu 2004, 2007a).

⁶ Štekauer et al. (2012: 201) suggest counter-examples to this universal as well, however it seems that the universal is misinterpreted. They say "... if a language makes use of infixation, it may also be expected to employ prefixation and/or suffixation in word-formation. ... Exceptions to this assumption include Yoruba, which uses infixation but not suffixation, and Tatar, which uses infixation but not prefixation." Neither language is a true exception to the statement with "and/or," since Yoruba has many prefixes, and Tatar has many suffixes.

Furthermore, in each language the claim that infixes exist seems unfounded. Yoruba has a reduplicated construction where /-ki-/ is inserted between reduplicated stems (ilé 'house' ilé -ki-ilé 'any house'), and though /-ki-/ is referred to as an "infix" (Bamgboṣe 1966: 153), it is clearly an interfix—an affix placed between two stems, which is semantically empty. The Štekauer et al. (2012: 201) reference to a Tatar infix (more specifically "-t-" on p. 202) does not correspond with standard morphological analyses of this language, which includes a range of mono-consonantal interfixes or linking elements (see Section 9.1.1) appearing between root and stem-forming suffixes none of which are /-t-/ (Ganiev 2006: 140).

Another newly described language that seems to counter-exemplify this universal is Kri, a Vietic language within the greater Austroasiatic/Mon-Khmer family (Enfield and Diffloth 2009).⁷ Though Proto-Mon-Khmer had derivational prefixes and infixes (see Section 9.5), sound changes have rendered prefixal derivations opaque in Kri, limited to a few scattered word pairs with similar initial consonants. However, inherited infixal derivational relations remain transparent: *kooq* ‘to live,’ *krnooq* ‘a house’; *keep* ‘to pinch,’ *krneep* ‘tongs, pincers’; *sat* ‘to get one’s foot stuck,’ *srnaat* ‘a foothold,’ etc. In fact, the only transparent derivational relations appear to be infixal! These include nominalizing <rn> just exemplified (infixated after the first C of a verb stem), as well as causative <a> and verbalizing <a> (infixated after CC- of CC-initial stems) (Enfield and Diffloth 2009: 44). Where segmentally specified bound morphemes are involved, then, morphological systems with infixation, but no prefixation or suffixation are possible.

As all the infixes in both Kri and Pingding Mandarin are derivational, and there are no other productive affixes, we also see that languages can have more derivational infixes than inflectional ones. In fact, it has been claimed that there is a strong cross-linguistic tendency for infixes to be derivational (Ultan 1975: 168f, Bybee 1985: 97, 110). Based on a survey of 70 infixing languages, Moravcsik (2000) puts the generalization in these terms:

...there is a broad tendency: the base-internal positioning of infixes tends to be iconically reflective of the fact that their meanings are closely tied to that of the base... *infixes are generally derivational, rather than inflectional* reflecting the closer semantic link between base and derivational affix than what holds between base and inflectional affix... (Moravcsik 2000: 548)

Is this true? Do derivational infixes far out-number inflectional infixes cross-linguistically? Clearly, a great deal depends on precisely how one classifies inflection vs. derivation, but even under a conservative approach, where only nominal agreement features of person/number/case and verbal tense/aspect/mood features are included in the category of inflection, languages with inflectional infixes seem to be just as common as those with derivational ones, and single languages can show a strong preference for inflectional over derivational infixes. Yu (2007a) includes a database of 154 infixation patterns from 112 languages representing 26 different phyla, with languages chosen on the basis of having any infix whatsoever. Of the 87 examples of non-reduplicative infixation included, 35, approximately 40%, are inflectional. In some languages, like Alabama, inflectional (person-marking) infixes outweigh derivational (comparative) ones 2:1, and in other languages, like Archi (Nakh-Daghestanian), all recorded infixes are inflectional. This same database of non-reduplicative infixation includes only two Papuan languages, Hua and Yagaria, both of the Trans-New Guinea phylum, and both with only inflectional infixes.

⁷ Kri also has reduplication which could be regarded as suffixing: *careew* ‘green,’ *careew-reew* ‘greenish,’ though this is not evident with monosyllabic bases like *tanq* ‘to chop up (meat),’ *tanq-tanq* ‘to chop up (meat) into tiny pieces.’ If reduplication is productive, and is included as a case of affixation as opposed to compounding, then Kri would not be considered a counter-example to Greenberg’s universal.

While one might argue that languages of New Guinea are under-represented in Yu's survey, as well as other studies of infixation, additional scouring of the literature has only turned up more instances of inflectional infixation: Eipo, and other Mek languages of Irian Jaya show inflectional tense/aspect and object-agreement marking infixes (Heeschen 1978); Au, a Torricelli language has verbal agreement infixes (Scorza 1985: 226); Yeri, another Torricelli language, has imperfective <m> and mirative <p> (Wilson 2011); and in Barupu of the Skou family, an inflectional <C> infix is part of the verbal person/gender/number marking system (Corris 2008). Manambu, a Ndu language, has a single derivational infix /-ka-/, intensive, used with non-agreeing adjectives (Aikhenvald 2008). At present, then, it appears that among the 800 or so non-Austronesian languages of New Guinea, at least 1% show infixation, and inflectional infixes out-number derivational ones by a wide margin.

While on the topic of areal tendencies, Moravcsik's (2000: 548) remark that "No infixes seem to have been reported from (non-Semitic) Africa and Australia" should be updated. Yu (2007a) includes seven Australian languages in his survey, all with infixing (internal) reduplication. To date, however, there are no known cases of non-reduplicative infixes in Australian Aboriginal languages.⁸

Within Africa, infixation is rare outside of the Afro-Asiatic family but attested.⁹ In both Birom and Noni, Niger-Congo languages, a noun-class infix <w> appears to have evolved via historical metathesis from earlier prefixes (Blevins and Garrett 1998). Overall, eight Niger-Congo languages are included in Yu's (2007a) survey, most with inflectional infixes. In addition, Bole (Gimba 2000), an Afro-Asiatic language, has a pluractional infix <ki> and Hadza, thought to be an isolate, also has a pluractional infix <kV> (Miller 2008).¹⁰

Another commonly held view of infixation, and, in particular, derivational infixation, is that it is less stable than other affixation types (cf. Ultan 1975: 185). The question of infix stability is taken up in the following section.

9.5 THE STABILITY OF INFIXES

A strange feature of popular writing on language is the common practice of referring to a modern spoken language as "ancient" or "one of the oldest languages on earth." In some cases, authors are clearly referring to a culture that appears to have existed with

⁸ The "verb-splitting" described by Henderson (2003) for Arrernte, a Pama-Nyungan language, inserts whole words, and even phrases, in the middle of other words, respecting prosodic, but not morphological boundaries. As outlined in Section 9.1, this process is designated as *tmesis*, since the inserted morpheme is free, not bound.

⁹ Within Afro-Asiatic, infixation is attested for Semitic, Cushitic, and Omotic languages.

¹⁰ See Section 9.5 on the stability of infixation over time, and Section 9.6 on the borrowability of infixes in intense contact situations. Both of these factors suggest that potentially cognate infixes make good starting points for hypotheses of genetic relatedness or extensive contact in pre-historic times.

little change for millennia, but in other cases, the claim that some languages are much older than others is clear:

The last speaker of an ancient tribal language has died in the Andaman Islands, breaking a 65,000-year link to one of the world's oldest cultures... Bo is one of the 10 Great Andamanese languages, which are thought to date back to pre-Neolithic human settlement of south-east Asia. (Watts 2010)

Since spoken languages are constantly changing, no modern language is entirely ancient, in the sense of reflecting precisely the same sound, word, and sentence structures as the language from which it descended, and Aka-Bo of the Andamans is no exception.¹¹ However, the field of historical linguistics certainly provides cases of words and morphemes whose form and meaning have remained relatively stable across time. And in these cases, modern words are indeed “ancient” as they have essentially the same properties as those used in pre-historic times.

One of the best examples of a language family with ancient words and morphemes is Austronesian (Greenhill et al. 2008, Blust 2009, Blust and Trussel 2010). Proto-Austronesian (PAN), the reconstructed mother language of more than 1,000 modern Austronesian languages, is thought to have been spoken approximately 6,000 years ago on the island of Formosa, present-day Taiwan. By use of the comparative method, aided by high quality data from hundreds of Austronesian languages, hundreds of lexical reconstructions are widely agreed upon. Remarkably, many of these proto-forms are reflected without change in modern languages as a consequence of stable sound patterns and cultural continuity. For example, the PAN verb **bilan* ‘to count, calculate; hold valuable,’ has many modern reflexes which appear to be nearly identical to the word as spoken 6,000 years ago: in Taiwan, Kavalan *biran* ‘to count’; in the Philippines, Bontok *bilan* ‘to count; the importance or worth of people’; and from the island of Flores in eastern Indonesia, Manggarai *bilan* ‘to count, calculate’ (Blust and Trussel, 2010).

Of interest to this study is the fact that two productive derivational infixes are reconstructed for PAN: **<um>*, a marker of actor focus, and **<in>*, a past/perfective marker and a marker of deverbal nouns.¹² Blust (2009: 370–88) details the history and synchronic status of both of these. Though the status of **<um>* as a derivational infix could be debated, since it serves a role similar to that of case and topic markers, there is little question that **<in>* had a derivational function in PAN, deriving nouns from verbs, and that the form and function of this infix has been maintained in the majority of languages in which it was directly inherited. Among the Formosan languages, we find Atayal *<in>*, marker of deverbal nouns, and the same deverbalizing infix is found in different subgroups, for example Toba Batak (Western Malayo-Polynesian), Wetan (Central Malayo-Polynesian), and Raluana (Oceanic) (Blust and Trussel 2010).

¹¹ In fact, historical reconstruction of this family has only just begun (Blevins forthcoming b).

¹² A plural infix /-ar-/ is also reconstructed but has far fewer reflexes in modern languages. See Blust (2009: 377–80).

The ancient status of PAN *<in>, a derivational infix, allows us to evaluate a widely held view about infixes: that they are unstable and short-lived. This view was first put forth by Ultan (1975: 185), and later repeated by Moravcsik (2000: 549), who associates the historical devolution of infixes with fossilization and/or externalization of infixes as prefixes or suffixes. However, within Austronesian, the picture is not one of instability, but of stability. More than 200 languages show reflexes of PAN *<in>, and in the great majority of these (approximately 75%), the morpheme remains an infix.¹³ Though some languages have lost this infix (e.g. Puyuma), the same language has a reflex of PAN *<um>, suggesting that the loss of the nominalizing infix is not a consequence of infix instability. Further, while the same language, Puyuma, does show fossilized instances of <in>, as expected with highly lexicalized forms (e.g. PAN *C<in>apa ‘smoked meat or fish’ < *Capa ‘to smoke meat or fish,’ Puyuma T-in-apa ‘what is grilled or roasted; smoked millet,’ from Blust and Trussel (2010)), externalization of *<in> as a prefix or suffix is not found, suggesting that where a derivational infix is lost, its loss may be no different from a range of other bound and free morphemes which simply fall out of use. In the case of PAN *<in>, a factor contributing to loss may be the existence of other deverbal/nominalizing morphology.

If Austronesian is representative of languages with productive derivational infixation, then, based on a well known history of 6,000 years, we can conclude that a derivational infix with high functional load is stable in terms of its morphological form as an infix, and in terms of its derivational function, as a deverbalizer.¹⁴

Austro-Asiatic, which includes Mon-Khmer and Munda languages, also appears to show modern reflexes of an ancient system of infixation (Shorto et al. 2006). However, there is less agreement among specialists in this area as to the nature of proto-Austro-Asiatic reconstructions, the place of the homeland, the approximate age of the proto-language, and the internal subgrouping of the family. Nevertheless, these infixes also show great stability over time, and, as discussed in the following section, have been the target of borrowing from neighboring unrelated languages.

Following Sidwell (2008: 257–64), Proto-Mon-Khmer has at least three productive derivational infixes, all nominalizing: *<n>; *<m> (agentive); and *<p> (instrumental). If Proto-Mon-Khmer and Proto-Munda diverged approximately five to six thousand years ago, then reflexes of these derivational infixes could be as old as the Austronesian infixes mentioned above. Though this is a smaller language family than Austronesian, with several hundred languages, the fact that the nominalizing /-n-/ infix is found in

¹³ In support of infixes as morphological adfixes, Plank (2007: 60) notes the ‘re-externalization’ of <um> in Tagalog as a prefix. However, of the 34 languages in Blust (2009: 383–4) with productive reflexes of PAN *<um> ‘actor focus,’ only Pazeh, Cebuano, and Makah Melanau show externalization of the infix. Further, in the case of Tagalog, contact with non-infixing languages like Spanish and English may play a role.

¹⁴ Recall that the same morpheme had an inflectional role in PAN as well, marking past tense or perfect aspect on verbs. This inflectional function has been lost in some languages that show a reflex of nominalizing *<in>.

most Austro-Asiatic languages (Diffloth and Zide 1992: 159) is consistent with the view of derivational infixes as stable morphemes.

9.6 BORROWED (DERIVATIONAL) INFIXES

Bound morphemes have often been claimed to be the least likely elements to be borrowed in a contact situation (Whitney 1881, Haugen 1950, Weinreich 1953, Van Hout and Muysken 1994). Nevertheless, a growing inventory of this type of borrowing is slowly being amassed along with ways of assessing the type of contact situation facilitating it (Sanchez 2005). As far as infixes are concerned, outside of specialist literature on problems in historical morphology of particular languages in Southeast Asia, very little has been written on the topic of infix borrowability.

Since, by definition, infixation is more complex than prefixation and suffixation in requiring a phonologically-defined locus for placement within a base, one might imagine that infixes are less often transmitted laterally via language contact than other affixes.¹⁵ A stronger position, that infixes are unborrowable altogether, has been taken, most recently by the anonymous author of a column entitled “Significant Activity in Linguistics” in the Summer, 1995 Issue 25 of *The Long Ranger* (formerly *The Mother Tongue* Newsletter of the Association for the Study of Language in Pre-History).¹⁶ In this short column, which discusses possible cognacy of the Proto-Austro-Asiatic and Proto-Austronesian derivational infixes discussed in Section 9.6, the author implies that borrowing of this kind is impossible:

What about borrowing?... We will offer a prize to the first person who can demonstrate the borrowing of a true infix between any languages of the world. If some of us think that the borrowing of pronouns is rare or non-existent, that is still inherently more likely than the case of the Austric infix. Who will take up my wager? Who will win? (Anon., *The Long Ranger*, 1995, Issue 25)

Well, it seems the prize should probably go to Franklin E. Huffman, whose 1986 paper, “Khmer Loanwords in Thai” makes a very strong case for the borrowing of Khmer infixes into Thai, with documented productivity in native Thai roots.¹⁷ As would be expected under Thomason and Kaufman’s (1988: 46) borrowing scale, language contact

¹⁵ The Optimality Theory analysis of infixes as morphological adfixes reviewed in Section 9.2 predicts that infixes cannot be borrowed, since they do not exist. Rather, under borrowing, a prefix or suffix is expected, with placement of that affix dependent on the differing phonological constraint-ranking of the target language. The data in this section, then, provide another argument for infixes as morphological primitives.

¹⁶ There is no attribution of this column, nor, as far as I can tell, mention of a newsletter editor in the online version or documentation at <<http://www.people.fas.harvard.edu/~witzel/aslip.html>>.

¹⁷ Huffman’s collected works, including his comparative lists and field notebooks are available via the SEALANG archive at: <<http://www.sealang.net/archives/huffman/>>.

between Khmer and Thai speakers was intense and long-lasting, extending from the 13th to 18th centuries, with Khmer culture dominant at the start, but Thai culture dominating in the later stages. Given this, identification of direction of borrowing is difficult, but Huffman uses morphology as a key: once Indic loans are eliminated, Thai is essentially a mono-syllabic isolating language. Khmer, on the other hand, has a wealth of derivational prefixes and infixes, including a nominalizing instrumental /-n-/, causative /-Vm-/, and abstract nominalizers /-VN-/ (with initial CC clusters), /-Vmn/ (with initial singleton C) (Huffman 1986: 200). On this basis, Huffman is able to identify loans like those in (9), where Thai appears to have borrowed an infixed Khmer form, along with its base.

(9) Infix borrowing, from Khmer into Thai (Huffman 1986: 201)

<i>Khmer</i>	<i>Thai</i>	
a. kaət	kəət	'to be born'
k<amn>aət	k<amn>əət	'birth'
b. traŋ	troŋ	'straight'
d<am>raŋ	d<am>roŋ	'to straighten'

However, demonstrating borrowings with infixes is not the same as showing that infixation as a derivational process has been incorporated into Thai grammar. To do this, it must be shown that the infixes have been extended to non-Khmer stems, or that the process has taken on distinct properties in Thai grammar.

Huffman (1986) provides highly suggestive evidence for both productivity in Thai, and distinct semantics. Productivity makes the task of finding true Khmer loans difficult: the assumption that infixed Thai words in (9) are loans "is complicated by the fact that Thai may have borrowed so many derivatives of this kind that it perceived this derivational process as a subsystem in Thai and infixed some basic Thai roots by analogy" (Huffman 1986: 201). And, since some Thai roots have also been borrowed back into Khmer, one must find some way of distinguishing Thai roots infixed in Khmer (and borrowed back into Thai) from Thai-internal cases of infixation. In this instance, differences in meaning between base and non-infixed forms in the two languages are probative: "The most common function of infixation in Khmer is the derivation of a disyllabic noun from a monosyllabic verb, while in Thai the derivative is typically a stylistic variant of the base verb, or a semantically specialized noun" (Huffman 1986: 201), as illustrated in (10).

(10) Infix borrowing, with Thai semantic innovation in bold (cf. Huffman 1986: 202)

<i>Khmer</i>		<i>Thai</i>	
a. qaoy	'to give'	?uay	'to bestow'
q<amn>aoy	'gift'	?<amn>uay	'to bestow (elegant)'
b. daə	'to walk'	dəən	'to walk'
d<amn>aə	'trip'	d<amn>əən	'to proceed (royal)'

c.	-----	sǎŋ	‘sound, voice’	
	s<amn>iəŋ	‘sound, voice’	s<ǎmn>iəŋ	‘pronunciation, accent’ ‘pronunciation’

Since the semantics of Thai infixation as a productive process appear to differ from that in Khmer, a Thai base with this pattern would be indicative of the productivity of infixation in Thai. The pairs in (10c) illustrate just this: *sǎŋ* “sound, voice” is a Thai root, showing both the derived form and meaning expected under productive Thai infixation. The derived Thai form has apparently been borrowed into Khmer, as shown by the semantics associated with it.¹⁸ This example of apparent nativization of derivational infixation is striking, not only in light of the rarity of infix borrowing, but also because Thai is historically an isolating language.

A second well documented case of infix borrowing is described by Thurgood (1999) for Proto-Chamic, the ancient Austronesian language associated with the Champa Kingdom, and known from inscriptions dating back to the 4th century. This example of infix borrowing is perhaps less striking than the Thai case, since infixation already existed as a derivational process in the target language. On arrival in coastal Vietnam approximately 2,000 years ago, Chamic people came in contact with speakers of Mon-Khmer languages, with clear effects on Proto-Chamic phonology, morphology, and lexicon. The classic proto-Austronesian disyllable was rendered iambic, with significant reduction of the first syllable, and explosion of vowel qualities in the second; consonant clusters evolved, along with new laryngealized consonants; and Mon-Khmer loans constituted as much as 10% of the Proto-Chamic lexicon. In the area of borrowed morphology, Thurgood’s reconstruction of Proto-Chamic includes the deverbal instrumental infix *<ən>, a clear instance of borrowing from neighboring Khmer languages (Thurgood 1999: 239). Though this infix appears to have fallen out of use in most modern languages, it is attested in inscriptions, and in transparent derivational relationships in some modern languages, for example Chru *phà* ‘to plane,’ *p<ən>hà* ‘a plane’ (Fuller 1977: 78).

In sum, the productivity of derivational infixation in Mon-Khmer languages has given rise to at least two clear instances where derivational infixes were borrowed into unrelated languages via intensive contact. Modern Thai reflects this exchange, while infixation of instrumental *<ən> in Proto-Chamic has lost its productivity in modern Chamic languages. Infixes can indeed be borrowed, and the two best supported cases of this in the linguistics literature involve derivational infixes whose lineage in Austro-Asiatic, as summarized above, is long and robustly attested.

¹⁸ Though Huffman (1986) states that the non-derived Thai root has also been borrowed, Khmer *sǎŋ* “sound, voice” has not been found in the SEALang lexical database (<<http://www.sealang.net/monkhmer/dictionary/>>).

CHAPTER 10

CONVERSION

SALVADOR VALERA

10.1 INTRODUCTION

A classic reference in the field (Dokulil 1968a: 215) places derivation by conversion at the crossroads of morphology, syntax, and lexical semantics. In this, it is like other derivational processes, but conversion raises problems of description which result from the specific conditions that apply in derivation by conversion and do not in derivation by affixation or in compounding. These conditions are word-class change and formal identity between the base and the derivative (Tournier 1985: 171). The first condition raises cross-linguistic questions, because it places conversion, as van Marle (1985: 123) explains, in the framework of a "...larger and more complicated... system": the system of word-classes. The second condition, formal identity, raises cross-linguistic questions, for example, whether it exists in morphologically different types of languages and, if so, in what form and to what extent, and also language-specific questions, like the role of stress shift in derivation by conversion in English. These and other problems are well known in some Indo-European languages, where conversion stands out as especially controversial.

This chapter deals with some of the main problems related to the cross-linguistic description of derivation by conversion. The chapter first reviews the description of conversion as lexical derivation governed by the two conditions mentioned above (Section 10.2). These two conditions are then discussed in separate sections as follows: formal identity (Section 10.3) and word-class change (Section 10.4). The last section is a test of the distribution of conversion over a sample of languages (Section 10.5).

10.2 CONVERSION AS LEXICAL DERIVATION

10.2.1 The Interpretation of Conversion

Canonical conversion involves substitution of a new inflectional paradigm, new syntactic properties, and a new categorial meaning (Dokulil 1968a: 225). The new

inflectional, syntactic, and semantic properties associated with formal identity between the base and the derivative justify the interpretation that the same form now is a different lexeme. From this point of view, canonical conversion is part of lexical derivation, even if its substantial differences with respect to other derivational processes have given rise to the interpretation that it is not (cf. van Marle 1985: 8–9, 84, 145, Don 1993, 2004, O’Grady and de Guzman 1996: 157, Anward 2001: 731; cf. Olsen 1990 on the place of conversion between concatenative and non-concatenative morphology, and Hockett 1994: 173 on conversion between addition and subtraction in word formation).

One school of thought views conversion as “zero-derivation” or “zero-affixation” (cf. Marchand 1969: 359 *et passim*, Kastovsky 1969, 1980: 213–17, 230, 1992a: 291, 300, Adams 1973: 13 *et passim*, Kiparsky 1982c, Lipka 1990: 2, 85–6, Payne 1997: 224–5). This approach describes conversion on analogy with affixation, mainly within the syntagma (determinant/determinatum) framework, in order to preserve as much structural coherence and homogeneity within the system as possible, and also to avoid the addition of a process that is different from all the others described in word formation. In this framework, the principle is that a morphosyntactic operation has taken place that is analogous to others that serve the same derivational function, except that this operation has no overt expression (cf. Sanders 1988: 155 *et passim*, Payne 1997: 8). Originally, this framework contrasts zero-derivation and conversion, such that the latter is used as a synonym for transposition in the sense presented in Section 10.2.2 (cf. Marchand 1969: 360). Later discussions of the concept of zero-derivation lay stress on the fact that the contrast between conversion and zero-derivation is less important than the fact that the process in question is a derivational process, and so whichever term is used becomes “... basically a metatheoretical-formal question” (Kastovsky 1997: 85–6).

Although both conversion and zero-derivation imply the existence of derivational morphology, and although the contrast between the two is not always established, each term entails differences.¹ A number of arguments against the zero-derivation approach have been raised for several languages, starting with the absence of analogues for the zero affix, the correspondence of one and the same unit (zero) with a number of gender and case specifications, the existence of a range of analogues that may be equivocal or contradictory, or the zero affixes “different behavior to their supposed explicit counterparts” (cf. Zandvoort 1961, Sanders 1988, Olsen 1990: 191 *et passim*, Lieber 1992, and Štekauer 1996: 23 *et passim*; for a review of positions for and against this use of zero, cf. Pennanen 1984: 84 *et passim* and Bauer and Huddleston 2002: 1641). Corpus evidence has prompted other questions, for example whether the affix should be a prefix or a suffix and, more importantly, whether it makes sense to postulate several different zero-affixation rules when the syntactic nature of the input does not alter the semantic rule used in each case (cf. Plag 1999: 223–4). Despite the differences implicit in each concept, conversion and zero-derivation have since been integrated within one and the same descriptive framework in several approaches. In this case, they are used to establish a

¹ For a review of these, cf. Lyons (1977, cited in Sanders 1988).

contrast between different processes that give rise to similar results, for example using the term “conversion” for certain cases of partial conversion² and the term “zero-derivation” for full conversion (cf. Dokulil 1968b: 56 et passim, cited in Pennanen 1984: 82), or using “zero-derivation” for the description of change within the so-called participles (cf. Lieber 1981, cited in Spencer 1991: 20).³

An alternative interpretation places conversion outside derivation and presents it as lexical creation, or as “coinage” of a new word (cf. Lieber 1992, 2004: 95, 2005). The principle is that conversion can be explained without resorting to morphological rules, specifically as a second introduction of an existing word within a different category in the lexicon of a language. This principle assumes that, just as new words can be entered in the language, existing words can be entered under a new category too. This approach is preferred in generative morphology, although arguments have been raised against it based on phonological constraints and morphological restrictions in denominal conversion in closely related languages such as Dutch, English, and German (cf. Olsen 1990 and Don 2005a).

A totally different interpretation to the ones mentioned in the paragraphs above disposes of the existence of a process, whether it is called “conversion,” “zero-derivation,” or “relisting,” by denying the derivational or lexical process in the first place and then the conditions of word-class change and formal identity. In this approach, the morphological, syntactic, and semantic features that correspond to one or the other word-class involved in one form become instantiated in the actual occurrences of each word. The essence of this approach is that there is no relationship between these nominal, verbal, and/or adjectival manifestations. Word-class is a semantic specification of one and the same lexical item which is unspecified as regards word-class and may take on behaviors that are then associated with different word-classes.⁴ This has been argued often of nouns and verbs, of verbs and adjectives, and of nouns and adjectives. In this approach, considerable differences may exist between words in that, whereas some are underspecified as regards word-class and may then be

² That is, in syntactic derivation, or when the word converts syntactically (it displays a new syntactic function) but not the inflection usually associated with that function (see Section 10.2.2).

³ Participles are involved in conversion as regards verbs, nouns, and adjectives (cf. Trnka 1969: 185, Olsen 1990: 195, and Haspelmath 1996 on several languages; cf. also Dalton-Puffer 1996b: 39, Payne 1997: 38, Beard 1998: 60–1, Spencer 2005, Fanego 2006, De Smet and Heyvaert 2011, and Corbett forthcoming, on gerunds and participles as mixed categories). The data attested in the Typological Data System (<[http://language-link.let.uu.nl/tds/main.html#home\[2\]](http://language-link.let.uu.nl/tds/main.html#home[2])>) confirm the lack of a clear separation between participles (cf. Section 10.5).

⁴ This hypothesis dates back at least to Whorf (1937/1956), and different formulations of the idea have been proposed since then, usually for the distinction noun/verb. Interestingly, this does not apply to all nouns and verbs: some words are considered to be nominal or verbal by definition. For a review of this hypothesis, cf. Lipka (1971). Cf. also the discussion in Magnusson (1954: 19–21), Trnka (1969: 183), Chomsky (1970), Bergenholtz and Mugdan (1979), Hockett (1994: 175 et passim), Marantz (1997), Farrell (2001), and Don (2005a: 10). Olsen (1990: 187 et passim) and Baker (2003) argue against this hypothesis based on the review of the structural possibilities and the actual realization of such word-class specifications.

specified or instantiated as members of various categories according to the context, others are not. This description, where no word-class change actually exists, has also been contested, for example, by Baker (2003) for specific word-classes like nouns and adjectives, and by Don in a series of chapters (2004, 2005a, b) on conversion in Dutch, English, and German which are commented on in Section 10.3 regarding directionality.

10.2.2 Conversion as Derivation

The widespread placement of conversion in derivation stresses its operation on the levels of form and content. Of these, the former has been considered to be subordinate to the latter (cf. Leitner 1974, cited in Pennanen 1984: 85). This is in line with the semantic approach to word-classes described in Section 10.3. Nonetheless, according to some descriptions, the interpretation that conversion means substitution of a new meaning along with a new form and function, whether the change in meaning is more or less important than the formal or the functional change, is enough for considering conversion as derivation. Of these new formal, functional, and semantic properties, the latter have been questioned more often than the rest, and the semantic change involved in conversion has not always been recognized (cf. Sanders 1988: 157). This position is best illustrated with the contradiction implicit in Sweet's (1891–8, I: 39) perception of semantic change in conversion in English: "... although conversion does not involve any alteration in the meaning of a word, yet the use of a word as a different part of speech leads to divergence in meaning." In canonical conversion, it is assumed today that the converted unit retains its lexical meaning and changes its categorial meaning. This explains both the relation and the contrast between base and derivative noted above and elsewhere in the literature (cf. Lipka 1990: 86 and Štekauer 1998: 11 *et passim* on the former, and Marchand 1963a: 176–7 on the latter).

This is also relevant for the separation of conversion from cases where words may display different inflectional properties, partially different syntactic behaviors and different nuances of meaning without any formal mark, but which do not entail word-class change. This has been considered derivation and has been described occasionally in the literature as secondary word-class conversion (cf. Givón 1993: 70–1 and Payne 1997: 25; cf. also Anward 2001: 732). However, application of the term "conversion" for these cases goes against one of the two main conditions of conversion, specifically word-class change. As conversion ultimately implies the existence of a base and a derived term too or, in other words, the formation of a new word, the concept of conversion is usually not applied to these cases.

A more important issue is the kind of derivation involved in conversion. Conversion formalizes lexical meanings as different categorial meanings or, as van Marle (1985: 131) calls them, word-class values. When formal and functional correlations follow this semantic change, there is a strong argument for derivation of a new lexeme. When there is not, either because the form does not follow, or the new function can be

realized by a range of categories, including the original category of the word in question (so there is no word-class change in principle), the interpretation of a new category is open to discussion.

In this respect, Bauer (2005b) shows how form, function, and meaning are involved in word-class identification and in word-class change. It is also shown in Bauer (2005b) that the evidence of word-class change at one of these three levels does not always have a correlate at the others and how the lack of this correlate moves the lexeme in question away from the prototype of one word-class into the common space with another word-class. A range of interpretations is then available, according to the parameters on which word-classes have been established. The main difficulty for a cross-linguistic review of conversion is that different languages may rely on different parameters for different categories or word-classes. As will be mentioned in this section and in Section 10.4, this remains a problem in cross-linguistic research on conversion.

In any case, the description of conversion as a word taking new inflection, new syntax, and new lexical meaning leads to the separation between syntactic and lexical derivation. According to this, a word can take on only the syntactic behavior (transposition) or also the change in the semantic category or categorial meaning (conversion). The difference between strictly syntactic and lexical processes is supposed to express itself in the productivity, semantic predictability, lexicalization, and morphological potential or the output of each (cf. Kuryłowicz 1936, Marchand 1966, 1969: 228–9, Anward 2001: 731–2; cf. Neef 1999: 219 for a comprehensive review).

Several interpretations have been made of the assumption of new syntactic functions by one and the same word. The usual approach separates syntactic transposition from conversion (cf. Dokulil 1968b: 57, Olsen 1990, Denison 2001: 126, Bauer and Huddleston 2002: 1642). Syntactic derivation has also been described as partial conversion, especially if the new word-class entails an inflectional paradigm that is different from the one of the base but it is not shown by the converted word (cf. Sweet 1891–8, I: 39–40, Zandvoort 1972: 265–6). If the theoretical framework lays emphasis on syntax, conversion is considered to exist also when the “same” or “[very] similar” meaning occurs as different syntactic functions, or when only syntactic derivation exists (cf. Hockett 1994: 172, O’Grady and de Guzman 1996: 157, Anward 2001: 731–2). Finally, these cases have also been classified along with canonical conversion (cf. Paul 1982: 298, 305, under the term “categori(c)al transference”).

A contrast can be established here between the classic example where an adjective appears to head a noun phrase without assumption of the nominal inflection (for example, *the poor*) vs. the correlate where nominal inflection is assumed (for example, *the hopefuls*). Some authors claim that in both cases ellipsis results in conversion to different degrees, because the original adjective has moved away from the prototype of its word-class (cf. Pennanen 1984: 81, Tournier 1985: 175, Anward 2001: 732). However, the importance of inflection may vary. If the word-classes involved do not differ inflectionally and the only difference is in their syntax and in their meaning, there is little argument to distinguish the two cases. This is what happens in formally identical adjectives and adverbs in Spanish and in conversion between adjectives and adverbs in English (even

if the picture becomes more complicated here for the influence of past derivational processes that are no longer visible due to historical leveling of endings; for a comprehensive review and an interpretation of Spanish adjectives and adverbs in other terms than conversion, cf. Hummel 2000). Similarly, where the two word-classes involved in cases like the above, adjective and noun, share a large part of the inflectional potential or the inflectional potential is identical (if the adjective is non-gradable, as in Spanish classifying adjectives, for example, *un americano* “an American”), there is considerable difficulty in telling adjective from noun and, therefore, in establishing (degrees of) conversion. As was advanced above, these differences are hard to capture in a description of conversion, unless specific degrees of relevance for formal, functional, and semantic criteria within each word-class are established, in one and the same language and cross-linguistically.

10.3 FORMAL IDENTITY

The second condition for conversion, formal identity, may fail to apply in different degrees and with different degrees of importance. Some formal changes, like stress shift between English nouns and verbs, are relatively specific within conversion, because they do not occur in all word-classes. The relevance of the change is not always clear as regards word-class membership and, therefore, these cases are usually regarded as “peripheral,” “marginal,” or “irregular” conversion.⁵ In general, formal limitations are important in frameworks where inflectional properties take priority in word-class identification and where, therefore, conversion is mainly morphological. In the approaches where form takes primacy over content, these cases have been described as peripheral conversion. If content is given primacy over form, they are conversion.

A complex case is that of nouns and verbs which show a formal difference, as in stem-based derivation, for example, in morphologically related nouns and verbs in German like *Antwort* vs. *antworten* or *Frage* vs. *fragen*. These cases have been included and excluded from conversion in the literature, the difference between one and the other position lying in the importance granted to the formal contrast: it may be argued that the formal dissimilarity limits itself to the stem taking the minimal possible inflectional mark imposed by its new word-class. In this case, the change in the word-class does not imply any other formal change and, as in conversion in English nouns and verbs, it implies a new syntactic function and a new categorial meaning (cf. Kastovsky 1969,

⁵ Few authors refer to these changes in terms other than conversion (cf. Hockett 1994: 173, Iacobini 2000). Word-class change with such formal change is also presented thus in manuals like Payne (1997: 36) under the terms “suprafixation” or “suprasegmental modification” and Jackson and Zé Amvela (2007: 87), or under the term “(internal) modification” in Manova and Dressler (2005: 67 et passim) and Carstairs-McCarthy (2006: 752–3).

Marchand 1969: 363–4, Lieber 1992: 157).⁶ The opposite position rules out conversion in such circumstances based on a strict interpretation of the condition of formal identity, and interprets the infinitive ending as a derivative mark (cf. Fleischer 1982: 314, cited in Olsen 1990: 189; cf. also Pennanen 1984: 80). The latter view excludes the concept of conversion from inflectional (synthetic) languages and restricts it to the canonical type, as in English.

Bauer's (2005b) review of a selection of infinitival constructions shows the range of combinations of formal, functional, and semantic features that can be found in different languages. It also shows that the difference between certain languages limits itself to a feature imposed by the new word-class: the inflectional mark of the verb, as in French (for example, *The fierce attacks* vs. *Les devoirs*, Bauer 2005b: 26). A similar contrast can be established in one and the same language if the direction noun to verb (for example, Sp. *camino* / *caminar*, Fr. *nappe* / *napper*) is compared with the direction verb to noun (Sp. *deber* / *deberes*, Fr. *savoir* / *saviors*).⁷ In Spanish, as in French and other Romance languages, denominal verbs display the infinitive ending (even if it no longer has inflectional value), while deverbal nouns may not take inflection signaling the word-class, because the noun does not always carry any such inflection by default. This contrast can be expressed as the difference between applying the change on the stem or on the word. A range of intermediate cases are possible, where the new denominal verb can take adverbs and verbal dependents, and where the new deverbal noun cannot take the typical dependents of nouns (cf. Bauer 2005b: 26 et passim). But if we compare the canonical cases, where the new denominal verb displays all the verbal properties and none of the nominal properties, and the new deverbal noun displays all the nominal properties and none of the verbal properties (the infinitive ending does not show contrast with other verbal forms), the only difference is the formal requirements of the new word-class. There must be specific constraints for each of the two cases, because one (the denominal verb) is considerably more frequent than the other. However, from an output-oriented point of view, the stem and the word undergo the same word-class change and both change their inflectional paradigms, their syntactic function and their categorial meaning, and retain their lexical meaning, as in English *attack*. Is it justified to consider these as different processes, because the resulting word-class of each direction case had a specific requirement?

These examples are also worth discussing because they show a pattern where directionality is relevant. In fact, directionality is an extremely challenging feature that is gaining importance in conversion. The two main approaches for the identification of directionality, diachronic and synchronic, have proved largely insufficient for languages like English.⁸ In the diachronic approach, this is due to the limited availability and

⁶ Conversion is apparently applied to stems (*glass* vs. *glaze*) in old grammars of English (cf. Priestley 1762: 141, cited in Sundby 1995: 108) under the term "transmutation."

⁷ Directionality is established here based on the synchronic criteria available in the literature, specifically semantic dependence, range of usage, semantic range, and semantic pattern.

⁸ For a review of the criteria, cf. Sanders (1988: 158–9 et passim), Levin and Rappaport Hovav (1991: 130) and, more recently, Bram (2011).

reliability of the evidence that can be used for the identification of the direction (written chronological records), to the uncertain and changeable nature of such evidence, and partly also to an interest in a fully synchronic description of the features of conversion (cf. Marchand 1963: 176–8 and Kastovsky 2000: 121–2). In the synchronic approach, it is because the criteria proposed rely on inconclusive evidence or on principles that are not fully justified (for example, the appropriateness to establish directionality based on synchronic evidence), or without exceptions (for example, the principle that the derived term will always have a lower frequency of use or a more limited range of senses than its base).⁹ To the best of my knowledge, alternative approaches, like Sanders' (1988: 172–3) markedness relations, have not been applied systematically in cross-linguistic research on conversion. The issue becomes even more complex if the possibility is allowed for directionality to be according to individual senses rather than to lexemes as wholes (Plank 2010).

The point here is that, after decades of stagnation, directionality is now a relevant argument in the separation of processes that may result in noun/verb formal identity. Don (2005b) has proven differences between denominal verbs and deverbal nouns in Dutch that suggest that, despite leading to similar results as conversion, a lexical and a syntactic process may be at play simultaneously in the creation of what appears above as central conversion. Arad (2003, cited in Don 2005b: 10) also argues that different types of verb formations identical to nouns exist, some of which are noun-based and some root-based. The distinction, based on the semantics of each case, may be accompanied by phonological features (for example, stress). The hypothesis is that different directionality evidence may signal coexisting processes which are intrinsically different but which result in the same profile of different word-class and formal identity in the same language (cf. Don 2005b; cf. also Bergenholtz and Mugdan 1979).

10.4 WORD-CLASS CHANGE

Worded in a variety of ways in the literature, word-class change is one of the two conditions for conversion to exist. It expresses the difference between base and derivative in conversion, whether the stress is laid on the different inflectional paradigm, on the new functional potential, or on the semantic contrast between a word of a word-class and a related one that belongs to another.¹⁰

The difficulty with determining the limits of conversion by the application of the notion of word-class is that the very notion of word-class is not always clear-cut.

⁹ For a review of these criteria, cf. Marchand (1963, 1964), Trnka (1969: 185), or Adams (1973: 38–42). Bergenholtz and Mugdan (1979) added word length and vowel mutation for directionality in German.

¹⁰ For a review, cf. Pennanen (1984: 85) and Neef (1999).

Not only is it not clear how many and which word-classes should be analysed in an optimal grammar for a given language or how this should be determined, but the word-classes themselves are rarely watertight. That is to say, not only may a single word-form be analysed as belonging to different word-classes in different constructions, but there are also occasions on which it might not be clear which of two competing word-classes a form should be taken to belong to. If we cannot tell which word-class a word-form belongs to on a particular occasion, we cannot determine whether it has undergone conversion, and this is a relatively frequent case.

In a prototype-based theory of word-classes, as is often used in cross-linguistic research, spaces between categories are an inherent feature of the theory and, thus, a significant degree of overlap is allowed between categories (cf. Sapir 1921: 118 *et passim*, cited in Lipka 1971: 212).¹¹ The prototype approach helps explain the behavior of words in terms of word-classes, but has the opposite effect for the identification of conversion, because it allows intermediate categories and the transition from one word-class to another is expressed as a gradient. The gradient or the overlap between categories is based on the occurrence of formal and functional properties, but in many languages there is little specification as to what those properties are and how they should be ranked: for example, should all the inflectional categories be given the same importance? Should they be given more importance than functional or semantic properties?

Research on word-classes has contributed the notion of patterns according to which word-classes extend their potential to that of other classes over time.¹² The definition of these patterns may help identify the areas where categories overlap and the relevance of the shared ground. Haspelmath's (1996) separation of a word's syntax into internal and external syntax and into lexeme word-class and word-form word-class may shed light on the extent to which a word takes on features of several word-classes at the same time. However, this is still to be developed further before a usable framework for the definition of conversion is available. As pointed out by Croft (2000: 90–1), what cross-linguistic research describes as cross-linguistic patterns of variation is prototypes, not the boundaries between the prototypes. In those boundaries, word-class membership becomes “a matter of degrees” (Crystal 1967: 50), but the degrees have not been defined.

Bauer (2005b) laid emphasis on the need for an accurate description of what it means to change category, and on the importance that this has for cross-linguistic research on conversion. Bauer (2005b) also showed that conversion does not always display evidence that may qualify as canonical conversion, and that different degrees of conversion have to be allowed in different languages according to different criteria, because not all

¹¹ For an overview of the notion of prototype, cf. van Marle (1985: 132), Givón (1993: 51–3), Payne (1997: 7, 32, 37–8), and Saeed (1997: 37 *et passim*). For an alternative view, cf. Baker (2003).

¹² For example, attributive, predicative, adverbial, cf. Anward (2001: 730–1). On the notion of prototype and the diversification of the properties of central members of categories, cf., for example, Kemmer (1992: 145–7) and Anward (2001: 733–734). Cf. also Trnka's (1969: 184) notion of new, wider, or just different semantic boundaries of word-classes as a result of conversion and van Marle's (1985: 140) description of the inner structure of word-classes.

cases of conversion are canonical and because the limits between classes are a gradient. This is an important point, because it moves away from a position where one and the same concept of conversion is applied to different languages, to a position where different types of conversion that may apply in different degrees can be used for a unified cross-linguistic description of conversion.

Typological research has proposed a system of word-classes based on semantic categories that are then mapped onto phonological, morphological, and syntactic properties (cf. Anward 2001: 726).¹³ This is not far from the conventional approach, where the grammar of most languages is described based on the word-class system inherited from the classical tradition, even if, as Leisi (1985: 15) points out, “[t]he categories for our perception of the world are only created by individual languages, as classes of denotata”.¹⁴ The list of categorial meanings and their correspondence with word-classes is a matter of discussion, even if the literature seems to agree on the categorial meaning of certain classes, specifically nouns, verbs, and adjectives (cf. Croft 1984: 53 et passim, van Marle 1985: 144–5, Olsen 1990: 188, Miller and Fellbaum 1991: 204 et passim, Givón 1993: 53 et passim, Payne 1997: 32 et passim, Plag 1999: 220, Anward 2001: 726–727, Farrell 2001, Spencer 2005: 102 et passim, and Corbett forthcoming).

In conversion, if word-classes are established semantically and then paired with phonological, morphological, and syntactic properties, the chance of finding canonical cases of conversion cross-linguistically is reduced to the unlikely case where change of categorial meaning and its phonological, morphological, and syntactic counterparts are paralleled to a sufficient degree of coincidence in other languages too. Besides these cases, there are changes of categorial meaning that are not paired by all the formal and functional properties which are associated with the new category. The latter types of cases do not count as canonical conversion. Otherwise, conversion would be reduced to change of categorial meaning and we would be considering something different from conversion, and a different (exclusively semantically-based) descriptive framework, because word-classes as are usually known are a mixture of the formal, the lexical, and syntactic.

The degree and type of mismatches between the type of categorial meaning of each word-class and its associated formal and syntactic properties vary within one and the same language and across languages (cf. Anward 2001: 729–31 and Corbett forthcoming). For example, Don (2003) has shown how different but genetically close languages like Dutch and German diverge substantially as regards noun/verb conversion. Even if every semantic change of category always had systematic phonological, morphological, and syntactic manifestations, the comparison with the same categorial change and its manifestations in other languages would have problems finding parallels over a number of languages. This is because the formal or functional manifestations may not be the

¹³ Cf. however Trnka (1969: 183) for the claim that a semantic classification of words cannot have universal application. Cf. also Lipka (1971) on word-class identification based on syntax and morphology, without mention of meaning.

¹⁴ For a review of this interpretation of word classes, cf. Lipka (1990: 123 et passim).

same, may not have the same relevance for category membership in different languages, or may simply not exist across the languages considered.

In the same proposal on the typological definition of word-classes, it has been claimed that progress depends on transcending the conventional systems because these systems are language-specific generalizations that in typological research become cross-linguistic generalizations (cf. Anward 2001: 734). Extending language-specific generalizations cross-linguistically entails imposing the model of analysis of one or a group of languages to others (cf. Ansaldo et al. 2010). The very application of the concept of conversion to languages where it may not be as relevant as in some Indo-European languages, or may not be relevant at all, is another example of the same imposition.

As conversion depends on the concept of word-classes, progress in typological research on conversion may then depend on revision of language-specific manifestations of conversion that vary inter- and intra-linguistically. This means that the conditions set at the beginning of this chapter as standard requirements for conversion, that is, word-class change and formal identity, have to be interpreted differently according to the grammar of each language. The description of conversion depends on the appropriateness of the parts of speech model used for each language, and the criteria for word-class identification and the limits between word-classes vary between languages. Each language may have to rely on different criteria for identification of word-class change and, therefore, of conversion.

This goes against the standpoint mentioned above, in that, at least for conversion, language-specific word-class systems are needed, where word-class change really reflects what for each language matters as regards word-class change. The comparison of these results across languages would allow a much more accurate account of what conversion is in each language and what differences and similarities exist across languages, always according to the word-class system framework specific for each of the languages being compared.

Until such a descriptive framework is used, derivation by conversion can be compared across languages in the standard system of word-classes based on form, function, and meaning, and the standard definition of conversion has to be used. This, like the framework itself, is an idealization of what conversion should be in each language rather than what it is. The comparison resulting from the cross-linguistic use of the standard system of word-classes and of conversion will be a generalization, and its value will depend on how close the generalization is to the grammar of each language and, to some extent, to the theoretical standpoint on conversion preferred by the reader.

10.5 A CROSS-LINGUISTIC TEST

Cross-linguistic data analysis in word-formation is rare partly for some of the reasons outlined above: it relies largely on conceptual generalizations that are then applied on sets of data taken to be representative of a language for cross-linguistic generalizations.

Table 10.1 Different types of conversion and different types of languages. These results classify doubtful evidence as *Uncertain*

	Analytic	Synthetic	Polysynthetic	Mixed	Uncertain
Stem-based	1	10	1	0	2
Word-based	5	4	3	2	4
Stem- and word-based	0	8	0	2	2

When, as is the case, a range of views exists on what qualifies as the linguistic feature that is the object of cross-linguistic analysis and what not, every piece of data lends itself to different counting processes and results, all of which can later be interpreted according to a range of theoretical positions.

This test tries to minimize this relativism as far as possible, although at the cost of descriptive detail. This section is therefore a preliminary analysis of a sample of 64 languages tested for the variables language type, type of conversion (here polarized as stem-based or word-based conversion), and the word-classes involved in conversion (Table 10.1).¹⁵ The language sample and the data rely mainly on Štekauer (2008), later supported by specific references for classification and analysis of the data.¹⁶ The sample has the limitations inherent in the method used and allows only some general results (cf. Bauer 2010 on methods for cross-linguistic research, specifically on questionnaires by qualified informants). A larger sample size and perhaps a different data collection model would be necessary for more detailed results.

Four conclusions can be drawn from these tests. The first is that the probability for 44 out of 62 languages to use some kind of conversion differs significantly from what would be expected from a binomial (50/50) distribution (exact binomial test $p < 0.01$). This result has two readings: the first is that conversion is a widespread derivational strategy, at least if stem-based conversion qualifies as conversion; the second is that a considerable proportion of languages (even if it is only one third of the sample) do not use conversion in the wider sense. This proportion raises the question of why conversion is so widespread considering its low morphosemantic transparency¹⁷. Viewed from

¹⁵ The languages considered in the sample are, in alphabetical order: Afrikaans, Amele, Anejom, Bahasa Indonesia, Bardi, Belorussian, Breton, Catalan, Cirecire, Clallam, Czech, Dangaléat, Datooga, Diola Fogny, Dutch, English, Estonian, Finnish, French, Gã, Georgian, German, Greek, Hausa, Hebrew, Hindi, Hungarian, Ilocano, Italian, Japanese, Jaqaru, Kalkatungu, Karao, Ket, Konni, Kwakw'ala, Lakhota, Luo, Maipure, Mandarin Chinese, Māori, Marathi, Nelemwa, Portuguese, Romanian, Russian, Serbian-Croatian, Slavey, Slovak, Spanish, Swahili, Swedish, Tamil, Tatar, Telugu, Tibetan, Udihe, Ukrainian, Upper Necaxa Totonac, Vietnamese, West Greenlandic, Wichí, Yoruba, and Zulu.

¹⁶ The references are both general, for example, Haspelmath et al. (2005), and the Typological Database System, and specific (see the references at the end of the book).

¹⁷ According to Dressler (1987: 105, 116), only subtraction ranks lower.

the opposite end, that is, the languages where conversion is not recorded, it also raises the question of what replaces conversion in each language.

The second conclusion is that no significant differences are found in the frequency of occurrence of the different types of conversion with respect to a random distribution (chi square goodness of fit, $\chi^2 = 1.3$, $df = 2$, $p = 0.5$). In the languages where some kind of conversion exists, stem-based (14 languages), word-based (18 languages), and both (12 languages) occur with similar frequencies.

Concerning language type, the languages were classified according to their main tendency towards an analytic, synthetic, polysynthetic, or mixed model, even if this is admittedly an abstraction of their morphological properties. The low sample size prevents statistical analysis of the distribution of conversion and of conversion types according to the typological classification of the languages of the sample. As in the former test, the only type sufficiently represented in the sample (synthetic, with 22 languages) does not show any significant differences in the frequency of occurrence of stem-based, word-based-conversion, or both (chi square goodness of fit, $\chi^2 = 2.54$, $df = 2$, $p = 0.28$).

Regarding the word-classes involved, this test only considers the categories noun, verb, and adjective, because these are the word-classes whose existence has been demonstrated in cross-linguistic research (cf. Croft 2000: 85–6 and Baker 2003: 88–94, 169, 188–9, 238–9, 248–9, 263). It is, then, not by chance that, with a low sample size as the one used here, only analysis of the following word-class relation by conversion is allowed: adjective and noun, adjective and verb, and noun and verb, with respect to stem-, word-based conversion, or both. In this test, a significant association can be found between word-classes and type of conversion (Fisher's exact test, $p = 0.035$).

The analysis of the deviation between the expected and the observed frequency shows four major results (Table 10.2):

- 1 A higher frequency than expected in adjective and noun word-based conversion.
- 2 A lower frequency than expected in adjective and noun stem-based conversion.
- 3 A higher frequency than expected in adjective and verb stem-based conversion.
- 4 A higher frequency than expected in noun and verb by stem- and word-based conversion.

These results confirm structural tendencies described in the literature. Also, these are word-classes described as language universals, so it is logical that transfer between them may exist, sometimes in the form of what is understood as conversion, according to the results in 1 through 4 above. Nouns and verbs have been described as particularly similar or close,¹⁸ and conversion between them is one of the most frequent types among

¹⁸ Pennanen (1984: 88) cites Biese (1941: 339 et passim) on the long-standing closeness between nouns and verbs in Indo-European languages that still manifests itself as “root-nominalization.”

Table 10.2 Word classes and different types of conversion. Only the types where a sufficient representation are shown. These results classify doubtful evidence as *Uncertain*

	Uncertain	Stem-based	Word-based	Stem- and word-based
Adjective/noun	1	2	12	1
Adjective/verb	0	6	4	0
Noun/verb	0	15	14	7

all the structural possibilities that the different combinations of word-classes allow.¹⁹ Conversion involving other word-classes described in the sample is possible, but those word-classes are not shown here.

These results are generalizations about a canonical profile that hides nuances in one or the other language which are not captured by this framework, because the framework is not adapted to the structure of word-classes of each language, and to what conversion may mean in each language.

10.6 CONCLUSION

Dokulil (1968b: 58) argued that even such closely related languages as English and German may have to account for different possibilities of word-class transfer. Neef (1999) compared constraints between conversion in English and German, and again found substantial differences. Don (2005a: 14) argued for a unified analysis of conversion rather than for specific concepts of conversion in different languages, but such an analysis may not be possible in view of the considerable differences pointed out even for typologically close languages. Although empirical evidence may suggest the possibility of a unified analysis, it also suggests the existence of morphological restrictions that may not apply cross-linguistically. Finally, Bauer (2005b) laid emphasis on the need for specific criteria for the identification of what it means to belong to a word-class and what it means to change some of the properties of each word-class.

If the separation between word-classes is a gradient, conversion then may have to be viewed in terms of a gradient too and as one that belongs to the group of processes

¹⁹ Olsen (1990) reviews the structural possibilities between open word-classes and lays emphasis on the fact that the patterns that are actually realized limit themselves to conversion between nouns and verbs (in both directions) and then adjectives to verbs in English and German. Adams (1973: 42 et passim) also contains a review of these structural possibilities, and Kerleroux (1999: 93) reports similar findings for French.

that result in unmarked word-class change. Prototypical cases are central instances of conversion and differ from the rest according to the approach to conversion as an ideal abstract process.

In cross-linguistic research, as in research on one and the same language, the question is probably not just whether there is conversion or not, but how it should be defined for each language and which types there are. The accuracy of what is described as conversion will be as precise as the description of the word-classes for each language and how those languages represent the formal, functional, and semantic features of each word-class in the system of each language.

ACKNOWLEDGMENTS

I would like to thank Professors Laurie Bauer and Heinz Giegerich for their extremely helpful comments on earlier versions of this paper. Their contribution cannot be acknowledged enough by a mere note. Responsibility for all errors remains exclusively mine. I would also like to briefly pay tribute to Professor Dieter Kastovsky, one of the most renowned scholars ever in the study of zero-derivation, who passed away in autumn 2012. This work was funded by the Spanish Ministry of Economy and Competitiveness (research project FFI2012-39688).

CHAPTER 11

NON-CONCATENATIVE DERIVATION

Reduplication

SHARON INKELAS

11.1 INTRODUCTION

REDUPLICATION is a very widespread construction in the world's languages. Of the 368 languages in Rubino's survey of reduplication for the World Atlas of Language Structures online (<<http://www.wals.info>>, as of January 2013), 85% show some form of productive reduplication. Of the languages exhibiting productive reduplication, 277 (89%) exhibit both full and partial reduplication. Only thirty-five languages show only full reduplication, and none show only partial reduplication.

This chapter will survey reduplication with particular attention to its role in derivational morphology, drawing attention to the many important questions that remain open about the relationship between form and function in reduplication.

11.2 TYPES OF REDUPLICATION

Reduplication, broadly defined as the repetition of part or all of one linguistic constituent to form a new constituent with a different function, occurs at many points on a spectrum from phonologically defined partial reduplication to the repetition of syntactic phrases. The following canonical subtypes have been identified in the literature. The first three constitute the conventional spectrum of reduplication; this spectrum has been extended by Inkelas and Zoll (2005) to include the last two items as well:

- (1) a. Partial reduplication (repetition of a phonologically defined subconstituent of a word, with potential concomitant phonological modifications).

- b. Total reduplication (repetition of a root, stem or word, with minimal or no phonological modifications).
- c. Echo reduplication (reduplication of a word, with replacement of the onset or, sometimes, vocalism or internal material in one copy).
- d. Synonym reduplication (juxtaposition of two roots, stems, or words which are synonymous, antonymous, or closely semantically related).
- e. Syntactic doubling (in which a single word or constituent is mandated to occur twice in the same syntactic construction, often with obligatory intervening material).

Examples of partial and total reduplication are well known, surveyed, and analyzed from a theoretical perspective in such works as Moravcsik 1978; Marantz 1982; McCarthy and Prince 1986, 1995; Steriade 1988; Inkelas and Zoll 2005; Stolz et al. 2011; and many others. Since the purpose of this chapter is to explore connections between the subtypes of reduplication in (1) and their semantic or syntactic functions across languages, we will not dwell on the phonological properties of partial reduplication, although these are the focus of most of the literature of this type of reduplication. Suffice it to say here that partial reduplication falls into two general categories: reduplication of a minimal word (or metrical foot), as in (2a), and reduplication of a smaller constituent (e.g. a heavy syllable, or a light syllable) (2b, c):

- (2) a. Reduplicant is a minimal word: Diyari (multiple functions; Poser 1990: 132, citing Austin 1981; see also McCarthy and Prince 1996):
- | | | |
|----------|----------------|----------------|
| wila | wila-wila | 'woman' |
| kanku | kanku-kanku | 'boy' |
| ku kuna | ku ku-ku kuna | 'to jump' |
| tɪlparku | tɪlpa-tɪlparku | 'bird species' |
| ŋankaŋti | ŋanka-ŋankaŋti | 'catfish (pl)' |
- b. Reduplicant is a heavy syllable: Mokilese progressive reduplication (Harrison 1973, 1976, Blevins 1996: 523):
- | | | |
|--------|------------|------------------|
| pɔdɔk | pɔd-pɔdɔk | 'plant/planting' |
| soorɔk | soo-soorɔk | 'tear/tearing' |
| diar | dii-diar | 'find/finding' |
| andip | and-andip | 'spit/spitting' |
- c. Reduplicant is a light syllable: Tohono O'Odham pluralizing reduplication (Fitzgerald 2001: 942, 945):
- | | | |
|----------|----------------------------|-----------------------|
| pado | pa-pado | 'duck/ducks' |
| tablo | ta-tablo | 'shawl/shawls' |
| siminjul | si-siminjul (→ sisiminjul) | 'cemetery/cemeteries' |

Total reduplication, by definition, duplicates its morphological target without attention to phonological size or shape, as illustrated in (3):¹

- (3) a. Indonesian (Cohn 1989: 185)—Total reduplication of nouns encodes pluralization or diversity for nouns:
- | | | |
|---------------|-----------------------------|---------------------|
| búku | búku-búku | ‘book/books’ |
| minúman | minúman-minúman | ‘drink/drinks’ |
| kəməʃarakátan | kəməʃarakátan-kəməʃarakátan | ‘society/societies’ |
- b. Acehnese (Western Malayo-Polynesian, Sundic; Durie 1985: 39–40)—reduplication expresses emphasis:
- | | | |
|-------|-------------|---------------------|
| tambô | tambô-tambô | ‘drum (emphatic)’ |
| ma | ma-ma | ‘mother (emphatic)’ |

“Echo”-reduplication is a term often applied to total reduplication constructions in which the beginning of the second copy is replaced by a fixed substring (see e.g. Marchand 1969, Alderete et al. 1999). Familiar English examples include the ironic or pejorative Yiddish-derived pattern in which the fixed substring [ʃm] stands in as the onset of the copy, replacing any existing initial consonant(s) (see e.g. McCarthy and Prince 1986, Nevins and Vaux 2003) (4a). Kolami (Central Dravidian; Emeneau 1955) exhibits an areally common “et cetera” construction in which *gi* stands in for the initial (C)V of the copy (4b):

- (4) a. English shm-reduplication (data from Nevins and Vaux 2003):
- Pedro Schmedro
Iris Schmiris
breakfast schmeakfast
- b. Kolami echo reduplication (Emeneau 1955):
- | | | |
|--------|---------------|-----------------------------|
| maasur | maasur-giisur | ‘men (and the like)’ |
| kota | kota-gita | ‘bring it (if you want to)’ |
| iir | iir-giir | ‘water (and the like)’ |

Moving further down the spectrum in (1), synonym reduplication has been posited, particularly in the Morphological Doubling Theory of Inkelas and Zoll (2005), as a relative of reduplication. (See also Singh 2005, Peterson and Maas 2009.) In Morphological Doubling Theory, total and even partial reduplication arise from the double use of the same constituent, modified phonologically in the case of partial reduplication. Broadening the definition of “same” to mean “semantically identical” yields synonym constructions of the type discussed in Singh (1982) and illustrated in (5) for Modern Hindi. This construction pairs synonymous adjectives, the first of native origin and the

¹ There are some exceptions; for example, Kinande verb stem reduplication is total, but requires the reduplicant to be minimally disyllabic. See discussion in Section 11.4.2.

second of Perso-Arabic origin, to give an overall meaning of “[noun] et cetera”. Data are from Singh (2005: 271):

(5) a.	tan	badan	tan-badan	[Hindi]
	‘body’ [+native]	‘body’ [-native]	‘body, etc.’	
b.	vivaah	shaadi	vivaah-shaadi	
	‘marriage’ [+native]	‘marriage’ [-native]	‘marriage, etc.’	

Syntactic doubling is the term used here for constructions in which the same word is deployed twice in a grammatical construction, sometimes separated by linking material or other syntactic elements (see e.g. Ghomeshi et al. 2004 and Lehmann 2005), often with a fixed constructional meaning associated with the doubling. In Ewe, for example, Ameka (1999: 75, 96–100) describes a “deprecatory duplication” construction in which the two copies of a noun are separated by an adposition:

(6) a.	É-nyé	ɲútsu	gbó	ɲútsũ
	3SG-be	man	vicinity	man:HTS
	‘He is not a real man’ or ‘He is an effeminate/emasculated man’			
b.	tó	gbó	tɔ-é	ke kúme
	father	vicinity	father-DIM	this kind
	‘this kind of pseudofather’			

Such constructions often resemble synonym constructions in that one copy has additional modifications. In Ewe, the second noun in deprecatory constructions may exhibit a diminutive suffix (6b); in addition, the entire construction is marked, formally on the second noun, by a High tone suffix associated with nominal compounds (Ameka 1999: 96–7). Its effects can be seen in both (6a) and (6b).

This concludes our extremely brief survey of the types of phenomena that are relevant to a discussion of the grammatical functions of reduplication. Outside even of this broad spectrum lie, on the phonological end, consonant gemination or vowel lengthening processes which perform morphological functions (see e.g. Anderson 1992, Spencer 1998); and, at the syntactic end, reiteration of words or syntactic phrases for stylistic purposes, which appears to be a very, very common pattern cross-linguistically.

11.3 Functions of Reduplication

Reduplication serves a wide variety of functions cross-linguistically and within individual languages. These functions range over the standard morphological functions of derivation and inflection. Reduplication can also serve as a phonological concomitant of affixation (Section 11.4.1) and even simply as a semantically contentless structural repair (Section 11.4.2). We will briefly look at examples of each type of function before

posing the interesting question of whether specific functions of reduplication tend to be correlated with the specific forms of reduplication that we surveyed in Section 11.2.

11.3.1 Reduplication as Inflection

Inflection is defined in textbooks and handbook articles as morphology which creates different forms of a word (see e.g. Corbett 1999, 2010, Booij 2006), vs. derivational morphology, which creates new words. The difference lies in meaning; derivation changes lexical meaning, while inflection preserves lexical meaning. Reduplication frequently occurs in both types of morphology.

As an inflectional device, reduplication is perhaps most commonly found as an exponent of plurality in nouns (as seen above in Tohono O’Odham and Indonesian) and of pluractionality, in verbs. In verbs, pluractional reduplication is associated with actor number as well as with event pluralization (see e.g. Cusic 1981, Wood 2007). Wood and Garrett (2001) cite Yurok verb stem reduplication as a case of event-internal pluralization (7a), in which the action itself takes place several times. The Mani (South Atlantic) example in (7c), derived from the unreduplicated counterpart in (7b), illustrates the use of verb reduplication to pluralize an argument (Childs 2011: 179):²

- (7) a. ckem ckem-ckem ‘to count/to make small tattoo marks’
 pegon peg-pegon ‘to split/to split in several places’
 prkwrh(s-) prkw-prkwrh(s-) ‘to peck or knock/to peck or knock repeatedly’
- b. pé fók sàkàtà nár gbèn
 PRO.HAB perform sacrifice cow tomorrow
 ‘A cow will be sacrificed tomorrow’
- c. pé fók-fók sàkàtà sì-nár gbèn
 PRO.HAB perform sacrifice NCM-cow tomorrow
 ‘Many cows are being sacrificed tomorrow’

Beyond pluralization, reduplication in nouns has occasionally been identified as marking other inflectional categories as well. This is less common in nouns than in verbs. Reduplication as case-marking is rare but has been documented in the closely related languages Chukchi (Dunn 1999) and Koryak (Bogoras 1969), where it marks absolutive case. Another uncommon but attested inflectional nominal function of reduplication is the encoding of possession categories. In Tarok (Benue-Congo), partial reduplication encodes third person singular possessive (Robinson 1976). In Arosi (Central-Eastern Oceanic; Lynch and Horoi 2002), partial reduplication marks 1st and 2nd person possessive constructions.

Turning to verbs, reduplication is very frequently used to encode a variety of aspectual distinctions. Clearly related to the event pluralizing functions mentioned above,

² The NCM prefix on “cow” is also pluralizing; without verb pluralization, the utterance would mean “Cows will be sacrificed tomorrow” (Childs 2011: 179).

the aspectual categories of frequentative, repetitive, continuation, and progressive are among the most common for verb reduplication cross-linguistically. As Rubino notes in his (2005) survey, other distinctions are commonly marked with reduplication as well. In Alabama (Muskogean), reduplication encodes imperfective aspect (Hardy and Montler 1988), while in Till (Salish), reduplication marks inceptives (Reichard 1959). Perfective reduplication in Indo-European is well-studied (e.g. Niepokuj 1997). Kiyomi (1995), in her survey of Malayo-Polynesian, cites examples of progressive and habitual reduplication, in addition to other less common aspectual distinctions.

Rubino (2005) also reports from his survey that reduplication can encode tense. Tagalog is often reported to mark the future tense via CV prefixing reduplication (see e.g. Stolz et al. 2011); however, Schachter and Otnes (1972) argue that the construction is better characterized as marked contemplated aspect (see discussion in Kroeger 1993: 15). Mina employs a syntactic doubling construction to mark past tense: *kán i kán zá* ‘cross 3PL cross EE = they crossed [the river]’ Frajzyngier et al. 2005: 188).

Reduplication does not seem to be used, commonly or perhaps at all, to encode person or gender in verbal agreement, just as it is rare to find reduplication marking case or gender in nouns.

11.3.2 Reduplication as Derivation

Reduplication is used to perform a very wide array of derivational functions. A construction qualifies as derivational if it changes part of speech, or argument structure, or if it otherwise alters the meaning of a word sufficiently to result in a new lexeme.

Reduplication can convert nouns to verbs and vice versa. In Aroma, for example, total reduplication derives verbs from nouns (*mega* ‘magic,’ *megamega* ‘to make magic’) and nouns from verbs (*vawao* ‘to decorate,’ *vawaovawao* ‘decoration’) (Craig 1980: 127). Kiyomi (1995: 1162) demonstrates that a wide range of mappings from the categories of Verb and Noun to those of Verb, Noun, Adjective, and Adverb are attested in reduplication in Malayo-Polynesian alone. In her survey of thirty Malayo-Polynesian languages, twenty-two exhibited part of speech-changing reduplication. Numbers record the number of mappings of each type in Kiyomi’s sample:

(8)

N→	V	Adj	Adv
	(5)	(5)	(9)
	Woleaian	Kiribatese	Hiligaynon
	<i>shal</i> ‘water’ →	<i>maunga</i> ‘mountain’ →	<i>túig</i> ‘year’ →
	<i>che-chal</i> ‘to water’	<i>maunga-unga</i> ‘mountainous’	<i>túig-túig</i> ‘yearly’

V→	N	Adj	Adv
	(3)	(4)	(2)
	Tigak	Fijian	Indonesian
	<i>giak</i> ‘send’ →	<i>garo</i> ‘lust for’	<i>kira</i> ‘guess’ →
	<i>gi-giak</i> ‘messenger’	<i>garo-garo</i> ‘lustful’	<i>kira-kira</i> ‘at a guess’

Within part-of-speech category, reduplication also frequently creates new lexemes. For example, reduplication can perform valence-changing operations on verbs. Kiyomi (1995) reports cases from her survey of Malayo-Polynesian in which verb reduplication is valence-reducing (but no cases of the reverse): from Paamese, Kiyomi cites an example in which reduplication of the transitive verb *lahi-e* ‘is carrying him’ produces the intransitive *lahi-lahi* ‘is occupied.’ As another example of valence-reducing reduplication, Palmer (2009) cites reduplication as a source of unergative verbs in Kokota, e.g. *manei n-e-ke dupa-nou* ‘he_{RL-3S}-PFV punch-1SGO I = he punched me’ but *manei n-e du-dupa bla* ‘he_{RL-3S} RD-punch_{3.MT}’ ‘he was just punching’ (Palmer 2009: 193).

Rubino (2004) mentions associative constructions as a possible output of noun reduplication, citing Newman’s (1944) data from Yawelmani (Penutian) as an example: *k’ɔhis* ‘buttocks’ > *k’ɔ-k’ɔhis* ‘one with large buttocks’.

11.3.3 Reduplication in the Fuzzy Area between Derivation and Inflection

Any survey that attempts to sort morphological constructions into the categories of derivation and inflection will inevitably contain a discussion of constructions that do not neatly fit into either category (see e.g. Bauer 1996, Blevins 2001, Booij 2006, among many others). This is true of reduplication as well. Some of the most common functions of reduplication fall into this nebulous area: diminutivization, attenuation, augmentation, intensification, quantification, and conveying a sense of distribution or lack of control. None of these features are required in agreement systems or are structurally assigned like case, thus none are canonically inflectional; and all of these processes affect meaning, as derivation does, but it is arguable whether they create new lexemes. Even pluralizing morphology can be ambiguous in this way. Number is involved in agreement systems and ought, therefore, to be a prototypical inflectional construction. However, the difference between a singular or plural actor can deeply affect event structure and interacts in that way with valence-changing morphology, itself prototypically derivational; on the derivational character of pluractionality, see e.g. Štekauer (2012: 31).

Example (9) contains instances of reduplication constructions performing some of the more commonly found functions in this in-between category. Many other examples of this kind can be found in surveys of reduplication (Key 1965, Moravcsik 1978, Kiyomi 1995, Rubino 2004):

- (9) a. Diminution (Lushootseed; Urbanczyk 2006: 180):
 ǰəsəd 'foot' > ǰi-ǰəsəd 'little foot'
 báč 'fall down' > bí-báč 'drop in from time to time'
- b. Attenuation/limitation (Alabama, from Hardy and Montler 1988b: 408, Rubino 2004: 19):
 kasatka 'cold' > *kássatka* 'cool'
 lamatki 'straight' > *lámmatki* 'pretty straight'
- c. Intensification (Bikol; Mattes 2006: 7, 10):
 gabos 'all' > *gabos-gabos* 'all (more than appropriate)'
 tumog 'wet' > *tumog-tumog* 'soaking wet'
- d. Distributivity (Gurubasave; Gowda 1975: 39, Rubino 2005: 21):
 asem 'three' > *asem-sem* 'three each'
 ténet 'seven' > *ténet-net* 'seven each'
- e. Quantification (Manambu; Aikhenvald 2008: section 4.55):
 bap 'moon' > *bap-a-bap* 'month after month'
 təp 'village' > *təp-a-təp* 'every village'
- f. Collectivity (Maltese; Stolz et al. 2011: 271):
 tarag 'stairs' > *tarag-tarag* 'flights of stairs'
- g. Out-of-control (Lushootseed; Urbanczyk 2006: 203):
 dʔáq 'fall' > dʔáq-aq 'totter, stagger'
 čǎǎ 'spit' > sčǎǎ-ǎǎ 'cracked to pieces'

As we will discuss in Section 11.5, these common functions exhibit considerable iconicity, a property which has been much discussed in connection with reduplication.

11.4 Reduplication without Semantic or Syntactic Function

No discussion of the form and function of reduplication would be complete without at least a brief mention of the fact that reduplication often occurs without making any clear semantic or syntactic contribution of its own. This takes place in at least two ways. First, reduplication often occurs as a concomitant of overt affixation (Section 11.4.1), raising the question of whether the reduplication itself, or the affix, or the construction in which they co-occur, is the locus of meaning. Second, reduplication can sometimes occur as an apparent repair to a structural templatic problem, usually but not necessarily phonological in nature (Section 11.4.2). In such cases there is simply no way to ascribe meaning to the reduplication process.

11.4.1 Reduplication as Concomitant of Affixation

Both full and partial reduplication are commonly found as part of a complex morphological construction which also features ordinary affixation. Such cases are of considerable interest to morphologists, as they disrupt the idealized one-to-one mapping between meaning and form (see e.g. Anderson 1992: ch. 3, Dressler 2005). In Roviana (Oceanic), for example, the derivation of instrumental or locational nouns from verbs is marked simultaneously by total reduplication and the nominalizing suffix *-ana*; *ha^mbo* ‘sit’ ~ *ha^mbo-ha^mbotu-ana* ‘chair,’ *hake* ‘perch’ ~ *hake-hake-ana* ‘chair,’ *hale* ‘climb’ ~ *hale-hale-ana* ‘steps, stairs’ (Corston-Oliver 2002: 469, 472). The reduplication co-occurring with *-ana* serves no distinct semantic function of its own. In Hausa (West Chadic), one class of nouns forms its plurals via CVC reduplication and suffixation of *-i*, as in *gútsùrè*: ‘small fragment,’ *gútsàtsàrí*: (< *gútsàr-tsàr-í*), *gáǎdàm* ‘dispute, argument,’ *gáǎdàndámí* (< *gáǎdàm-dám-í*) (Newman 2000: 451).

In Ditidaht (also known as Nitinaht; Southern Wakashan), about forty suffixes trigger reduplication on the stems they attach to (Stonham 1994). For example, the “resemble” suffix triggers CV root reduplication (Stonham 1994: 40). Roots are underlined; reduplication-triggering suffixes are co-indexed with the reduplicant:

- (10) a. $\lambda'i_i$ - $\lambda'ic$ -ak'uk_i ‘RED_i-white-resembles_i’ = ‘flour’
 b. tu:_i-tu: χ -ubq-ak'uk_i ‘RED_i-scare-plant-resembles_i’ = ‘looks like a spruce tree (= juniper-leaved hair moss)’
 c. pi:_i-pi: λ a:q-k'uk_i ‘RED_i-liver-resembles_i’ = ‘resembles liver (= yellow pond lily)’

In cases like these, reduplication can potentially be viewed as a morphophonological accompaniment to affixation, much like ablaut or other morphophonemic alternations which commonly apply to bases of affixation. Alternatively, reduplication that accompanies affixation can be analyzed in terms of what Aronoff (1994) and Blevins (2003) term “morphomic stems,” that is, semantically empty stem-forming constructions producing stems that certain affixes select for; this is the approach taken by Inkelas and Zoll (2005), who analyze reduplication in the examples above as a semantically empty morphological process whose purpose is to form stems of a particular type. In Ditidaht, for example, the “resembles” suffix selects for the stem of the type formed by CV reduplication. Supporting evidence for a stem type analysis is that, as observed by Stonham, if two co-occurring suffixes both select for a reduplicated stem type, reduplication occurs only once (Stonham 1994: 49). Reduplication converts a root to a stem of the appropriate type, to which both affixes attach. This is true, however, only if the suffixes also create stems of that same morphological type. Stonham (1994: 57) posits two morphological levels, or stem types, in Ditidaht. Reduplication can occur twice if triggered by affixes in both levels:

- (11) ka-ka_i-kawad-ata χ _i ‘DIST_{RED}-RED_i-killer_whale-hunt_i’ = ‘hunting killer whales here and there’

Semantically empty, concomitant reduplication accompanies derivational and aspectual affixation, or what Stonham (following Haas and Swadesh 1932) calls stem-level affixation, but independent Distributive reduplication patterns with the outer layer of word-level affixes which Stonham characterizes as inflectional (Stonham 1944: 57).

11.4.2 Reduplication as Repair

Sometimes reduplication is used as a repair for ill-formed phonological or templatic structures. In such cases, it is clear that no meaning or syntactic function can be ascribed to the reduplication itself.

In some cases reduplication appears to be motivated as a repair for phonological sub-minimality. One intriguing example comes from Kinande, in which a morphologically compelled process of reduplication can be triggered to apply an extra time, semantically vacuously, for purely phonological reasons. In Kinande, verb stem reduplication is total (e.g. *-huma* ‘beat’ → *-huma+huma* ‘beat a little here and there’), but requires the reduplicant to be disyllabic. As a result, reduplication of a monosyllabic verb stem will result in triplication: *-swa* ‘grind’ → *swa-swa+swa* ‘grind a little here and there’ (Mutaka and Hyman 1990, Hyman 2009: 184).

A second example occurs in Chukchi, which also recruits an independent morphological process of reduplication, without its semantic connotations, to serve a phonological purpose. As mentioned above, Chukchi encodes absolutive case (in singulars) by partial reduplication (12a). This same reduplication pattern is used, minus its absolutive singular connotation, in absolutive plurals, which are normally marked with the suffix *-t*. Exactly when the root is CVC in shape, the absolutive plural shows the same reduplication found throughout the absolutive singular (Krause 1980).

(12) Root shape:

	<i>Absolutive singular</i>	<i>Absolutive plural</i>	<i>gloss</i>
a. CVCV	nute-nut tala-tal	nute-t tala-t	‘earth’ ‘pounded meat’
b. CVC	čot-čot tam-tam	čot-čot-te tam-tam-ət	‘pillow’ ‘growth’

Krause (1980) and Kiparsky (1986) attribute the phonologically conditioned plural reduplication in Chukchi to a prosodic minimality condition on noun roots; to be inflected (e.g. with plural *-t*), a noun stem should be larger than CVC. Some languages might epenthesize a vowel in this situation; Chukchi appeals instead to semantically null reduplication.

Peterson and Maas (2009: 225, 227) discuss the interesting case of Kharia (Munda), in which reduplication applies to masdars (predicates without TAM and person marking) exactly when the masdar would otherwise be disyllabic:

- (13) a. Polysyllabic lexical base: masdar is homophonous with base:
karay ‘done; (the act of) doing; do (ACTIVE)’
ayo ‘mother; become a mother (MIDDLE); accept (s.o.) as a mother (ACTIVE)’
 b. Monosyllabic lexical base: masdar is reduplicated:
ter ‘give (ACTIVE)’; *ter-ter* ‘(the act of) given; given; gift’
yo ‘see’; *yo-yo* ‘(the act of) seeing’

Peterson and Maas attribute this process to a historical requirement that predicates be disyllabic; reduplication was a semantically vacuous repair of subminimal predicates. In the modern language, they argue, the process has become restricted to masdars, applying to monosyllables even when the masdar-containing predicate as a whole contains other syllables (e.g. *aw* ‘live’ > *aw-aw-te*, in which the oblique case marker *-te* does not obviate reduplication in the masdar *aw-aw*; Peterson and Maas 2009: 227). Regardless of whether it applies to entire predicates (historically) or just to masdars (synchronically), however, it seems clear that reduplication is still a repair for phonological subminimality of a morphosyntactic constituent in Kharia.

An interesting case of syntactic doubling as a structural repair occurs in Chechen (Nakh-Dagestanian, Nakh), motivated by the requirement that some syntactic element precede and host a rigidly second position clitic (Conathan and Good 2000; see also Peterson 2001 and Good 2006 on the closely related language Ingush). This case is also discussed in Inkelas and Zoll (2005). As shown in (14), from Conathan and Good (2000: 50), chained clauses are marked by an enclitic particle *’a* (= IPA [ʔa]), which immediately precedes the inflected, phrase-final, main verb. The enclitic must be preceded by another element in the same clause. Two types of constituent may occur before the verb (and enclitic particle) in the clause: an object (14a), or a deictic proclitic or pre-verb (14). If neither of these elements is present, then the obligatory pre-clitic position is filled by reduplicating the verb (14c).³

- (14) a. Cickuo, [ch’aara = ’a gina]_{VP}, ’i bu’u [Chechen]
 cat.ERG [fish = & see.PP]_{VP} 3S.ABS B.eat.PRS
 ‘The cat, having seen a fish, eats it’ <cf. p 173 (7)>
 b. Ahmada, [kiekhat jaaz = ’a dina]_{VP}, zhejna dueshu
 Ahmad.ERG [letter write = & D.do.PP]_{VP} book D.read.PRS
 ‘Ahmad, having written a letter, reads a book’
 c. Ahmad, [ʃa = ’a ʃiina]_{VP}, dʃa-vaghara
 Ahmad [stay.INF_{Red} = & stay.PP]_{VP} DX-V.go.WP
 ‘Ahmad stayed (for a while) and left’

³ In glosses and cited forms, “B,” “D,” “V” represent prefixes encoding the gender class of the absolutive argument.

The Chechen reduplicant occurs in infinitive form, while the main verb is inflected. Inflected verbs require a different form of the verb stem than that used in the infinitive; in some cases the stem allomorphy is clearly suppletive, e.g. *Dala* 'to give' vs. *lwo* 'gives,' or *Dagha* 'to go' vs. *Duedu* 'goes.' As Conathan and Good (2000: 54) observe, the result is that Chechen can exhibit suppletive allomorphy differences between base and reduplicant (e.g. *Dagha* 'a *Duedu*, based on "go").

11.5 SEMANTICS OF REDUPLICATION: ICONIC OR NOT ICONIC

Perhaps the most common topic in discussions of reduplication is iconicity. To what degree are reduplication constructions semantically iconic, in the sense of "more form, more meaning"? Many surveys of reduplicative semantics have been devoted to this question.

Key (1965), based on a survey of forty-seven (mostly Native American) languages, emphasizes the commonality of emphasis, plurality, and augmentation. The assumption that reduplication is associated with the semantic concept of "more" underlies discussions in the literature of the iconic connection between form and content (Lakoff and Johnson 1980, Haiman 1997).

Moravcsik's (1978) thirty-five-language sample showed variety beyond these iconic meanings. Although plurality and intensity were frequent in her corpus, she also found diminution to be frequent, and observed that reduplication covers a wide variety of meanings and that the meanings that can be associated with reduplication can also be associated with nonreduplicative morphology.

Based on a close study of the semantics of reduplication in just one language family, Malayo-Polynesian, Kiyomi (1995: 1148) concludes that "reduplication can function either iconically or noniconically." Plurality is a canonically iconic meanings of noun reduplication; repetition and continuation are the canonical iconic meaning of verb reduplication. Kiyomi (1995: 1149) also identifies intensification as a canonical iconic meaning of reduplication; "in noun reduplication, some property of the noun in question is intensified in its reduplicated form, and in verb reduplication, the degree of an action is intensified." Hohenhaus (2004) discusses reduplication/repetition constructions in English, such as "food-food" (meaning "real food"), which convey prototypicality.

Regier (1994) employs a similar strategy in attempting to bring some coherence to the bewildering array of reduplicative functions identified by Moravcsik. Figure 11.1 is a proposal by Regier for relating some of the more peripheral, less obviously iconic meanings of reduplication to the ostensibly central function of repetition.

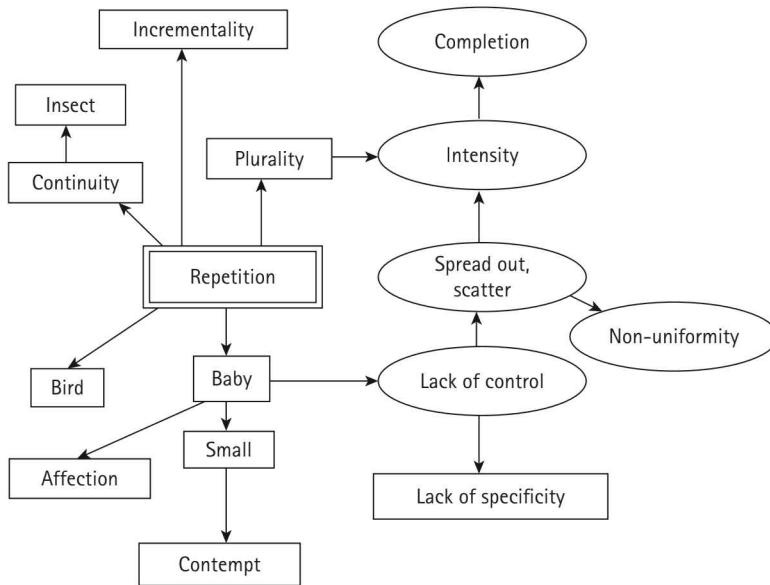


FIGURE 11.1 Radial category for the semantics of reduplication

Source: from Regier (1994).

Regier makes the point that the radial category centered around repetition is not specific, in its internal structure, to the formal morphological process of reduplication. Much of the semantic structure in Figure 11.1 is also found associated with nonreduplicative constructions; Regier calls particular attention to Slavic prefixes, including Russian *raz-*, whose meanings are represented by ovals in Figure 11.1.

Some meanings of reduplication, however, venture so far afield from the semantic categories related to ostensible core iconic meanings that notions of iconicity seem to lack usefulness altogether.

In some cases, this is illusory. For example, diminution, the apparent opposite of augmentation or intensification, is a common semantic correlate of reduplication cross-linguistically. This apparent contradiction in the senses common to reduplication is a topic of some consternation (e.g. Haiman 1997) and interest (e.g. Taylor 1992, Jurafsky 1996) in the literature. As Jurafsky (1996) makes clear, the nexus between augmentation/intensification and diminution is not unique to reduplication. Numerous nonreduplicative morphological constructions also have the diachronic or even synchronic property of expressing both seemingly contradictory properties. Jurafsky cites the example of *ahorita* ('now-DIM'), in which the suffix *-ita* produces the intensifying meaning of 'immediately, right now' in Mexican Spanish but the diminutivizing meaning of 'soon, in a little while' in Dominican Spanish (Jurafsky 1996: 534). Jurafsky proposes, along lines similar to Regier (1994), a radial category analysis of diminutive semantics which predicts the diachronic development of a range of possible meanings from an original meaning related to 'child' or 'small.'

But other cases of non-iconicity seem harder to argue away. Recall from Section 11.3.1 the cases of Tarok and Arosi, in which partial reduplication encodes possession.

Impressionistically, it appears that iconicity is most likely in total reduplication constructions, especially newer ones (as in creoles), and less likely in partial reduplication constructions. A thorough statistical survey of reduplication is needed in order to test the validity of this impression. The impression is related to another assumption commonly found in the literature, which is that total reduplication is the diachronic source of partial reduplication (see e.g. Bybee et al. 1994, Niepokuj 1997). If true, then the apparent iconicity cline would be a result of grammaticalization, showing semantic bleaching and drift and even reanalysis over time. However, this assumption is generally still untested by solid evidence, and some literature has expressed skepticism (Hurch and Mattes 2005, Stolz et al. 2011). In a detailed study of the natural history of verb reduplication in Bantu, a family exhibiting both total and partial reduplication, Hyman (2009) actually concludes that a likely scenario was somewhere in the middle for Bantu: an original scenario of root reduplication played out as total stem reduplication in some languages and as partial stem reduplication in others. Neither total nor partial verb stem reduplication represents the original state. See also Blust (1998) and Reid (2009) (among others) for discussion of the many pathways to CV partial reduplication in Austronesian. Unfortunately, reconstruction arguments at this level of detail are rare, and the origins of partial reduplication in the world's languages remain largely obscure.

11.6 AFFIX REDUPLICATION: REFLECTIONS ON ICONICITY

Affixes are frequently incidentally reduplicated as part of reduplication processes that target the stems they are part of. In some cases, however, reduplication targets individual affixes explicitly, as discussed in Inkelas and Zoll 2005. According to Roberts (1987, 1991), to express iterative aspect in Amele (Trans New Guinea, Madang), “the whole stem is normally reduplicated if the verb does not have an object marker, otherwise the object marker is reduplicated either in place of or in addition to the reduplication of the verb stem” (Roberts 1991: 130–1). Data are from Roberts (1987: 252–4) and Roberts (1991: 131):

(15) a.	qu-qu	‘hit’ (iterative)
	ji-ji	‘eat’ (iterative)
	budu-budu-e?	‘to thud repeatedly’
	ḡbatan-ḡbatan-e?	‘split-INF’ (iterative)
b.	hawa-du-du	‘ignore-3S-3S’ (iterative)
	gobil-du-du	‘stir-3S-3S = stir and stir it’
	guduc-du-du	‘run-3S-3S’ (iterative)
c.	bala-bala-du-d-e?	‘tear-3S-INF = to tear it repeatedly’

Van der Voort (2009) describes a case of person marker doubling in Kwazá, an isolate of the Brazilian Amazon, explicitly commenting that “[t]his kind of reduplication does not appear to be an iconic strategy, and it is not determined by the boundaries of phonotactic units like syllables, moras, or words but by morpheme boundaries” (Van der Voort 2009: 268). Verbs obligatorily inflect for subject person and optionally for object person. Past tense is not marked morphologically, but is expressed through the use of adverbs. However, remote past tense, in particular, is encoded by reduplicating person markers (Van der Voort 2009: 169), usually subject but in certain cases object markers. Compare (16a–b) to see the semantic effects of reduplication in (16b). A comparison of (16b–d) (Van der Voort 2009: 270–1) shows that the reduplicant copies the person marker regardless of phonological shape and size.

- (16) a. la'to o'ja-da-hĩ-ki ze'zeĩju-di-rjĩ
 yesterday go-1S-NOM-DEC Zezinho-POS-area
 ‘Yesterday I went to Zezinho’s place’
- b. ja o'ja-da-'dai-hĩ-ki txa'rwa oja-'he=(bwa)-da-ki
 already go-1S-1S-NOM-DEC first go-NEG=finish-1S-DEC
 ‘It has been a long time since I went there. I haven’t been there since’
- c. aure-lɛ-'nã-axa-axa-le-hĩ-ki
 marry-RECI-FUT-1P.EXCL-1P.EXCL-FRUST-NOM-DEC
 ‘We were going to marry (but we didn’t, long ago)’
- d. tsi'cwa-xaxa-xaxa-hĩ-'r Ba'hoso te'ja
 begin=2P-2P-NOM-INT Barroso side
 ‘Did you (plural) start (opening the trail) on the side of Barosso? (two years ago)’

Van der Voort argues on the basis of reduplicant shape that the construction is affix reduplication, not syllable or foot reduplication.

In Boumaa Fijian (Oceanic), stems formed by spontaneous or adversative prefixes mark plurality by individually reduplicating both the prefix and the root, in an AB → AABB pattern (Dixon 1988: 226):

- (17) ta-lo'i 'bent' ta-ta-lo'i-lo'i 'bent in many places' [Boumaa Fijian]
 ca-lidi 'explode' ca-ca-lidi-lidi 'many things explode'
 'a-musu 'broken' 'a-'a-musu-musu 'broken in many places'

As in Kwazá, the fact that the phonological size and shape of the Boumaa Fijian reduplicants varies with the size of the morpheme being reduplicated suggests strongly that this is morpheme doubling, not phonological copying motivated by the need to flesh out an abstract, phonologically skeletal morpheme.

In all three of these cases, the semantic content of the affix being reduplicated seems unrelated to the semantics of the reduplication construction.

(18)	Function of (unreduplicated) affix	Function of affix reduplication
Amele	object marker	iterative aspect
Kwazá	(subject) person marker	(remote) past tense
Boumaa Fijian	spontaneous, adversative	event plurality

Although the meanings of reduplication in these three examples are all iconic to a medium or high degree (iteration and plurality being central meanings of reduplication cross-linguistically), the semantic connection to the reduplicated affix seems quite arbitrary.

So far, cases of affix reduplication appear to fall into the domain of inflection, not derivation. This generalization needs to be tested further. If it holds up, it may suggest something important about the important relationship between form and function in reduplication.

11.7 FORM AND FUNCTION IN REDUPLICATION

An interesting question in the study of reduplicative function is whether form and meaning are correlated. Reduplicants come in a variety of sizes; reduplication performs a variety of functions, some highly iconic and some less so. Are the scales related at all?

A null hypothesis might be that total reduplication is associated with the more iconic end of the function scale, whereas partial reduplication is associated with a less iconic, more semantically diverse range of meanings. For example, in his study of Bantu verb stem reduplication, Hyman (2009) observes that it is only the smallest (syllable-sized) reduplication constructions in which habitual or imperfective aspectual meanings are found. Total verb stem reduplication in Bantu tends to have more transparent, characteristic functions such as attenuation or intensification. Given that reduplication in creoles tends to be total rather than partial, and given that reduplication in creoles tends to be more iconic than reduplication in languages with longer histories, such a correlation is likely to hold up statistically cross-linguistically as well, once a suitable survey is done. Echo reduplication, which tends to be associated with a smaller range of meanings, also tends very heavily to be total.

That said, there is still a large diversity of meanings to be observed within total reduplication, as demonstrated by the recent survey by Stolz et al. (2011). It is also the case that many partial reduplication constructions have meanings near the center of Regier's diagram of reduplicative semantics. Any conclusions will have to be statistical, not categorical.

It is also probably unwise to lump all partial reduplication together. For example, if one is pursuing the hypothesis that form and function are correlated, one might wish to distinguish between partial reduplication involving minimal words and partial

reduplication involving smaller (syllable-sized) constituents. It might, for example, turn out to be the case that the grammatical function of minimal word-sized partial reduplication constructions might more closely resemble that of total reduplication, vs. the partial reduplication of smaller constituents.

Within Generalized Template Theory (GTT; e.g. McCarthy and Prince 1994a, b, Downing 2006, Urbanczyk 2006), a different distinction within partial reduplication has been hypothesized: affix vs. root reduplication. According to the precepts of GTT, reduplicants are classified either as affixes or as roots. Note that this is not correlated with what part of the base is copied; it is a property just of the reduplicative morpheme itself. Downing (2006), working in the most advanced form of GTT, proposes that reduplicants assume the canonical shape of roots or affixes within the language. Thus in a language in which all roots are minimally bimoraic, root reduplicants must also be. If affixes in a language are maximally syllable-sized, affix reduplicants will also be. Urbanczyk appeals to the root/affix distinction to characterize two types of reduplication in Lushootseed. The proposed Diminutive reduplicant is CV in shape (with a reduced vowel), while the proposed Distributive reduplicant is CVC in shape (with a full vowel).

(19) Lushootseed reduplication:

a. Diminutives (reduplicant = type 'Affix')

'foot'	ǰəsəd	→ ǰi-ǰəsəd	'little foot'
'animal hide'	s-kʷəbšəd	→ s-kʷi-kʷəbšəd	'small hide'

b. Distributives (reduplicant = type 'Root')

'foot'	ǰəsəd	→ ǰəs-ǰəsəd	'feet'
'bear'	s-čətʰəd	→ s-čət-čətʰəd	'bears'

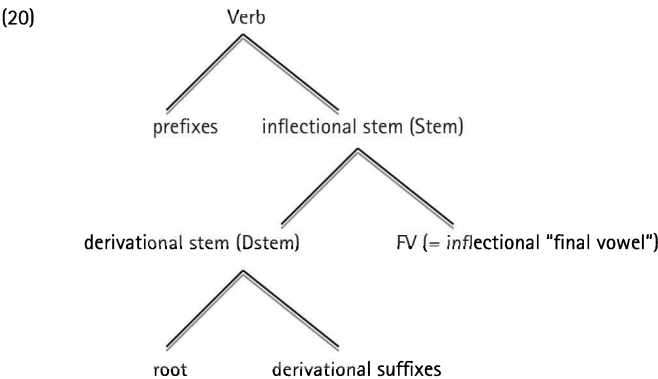
Urbanczyk (2006) attributes the phonological shapes of the two types of reduplicant to their classification as Affix (constrained to be as small a syllable as possible) and Root (constrained to be minimally bimoraic).

In this particular case, both diminutive and distributive meanings for reduplication are quite common, and it is hard to call either one more central. However, a profitable future research program might search for statistical tendencies in the cross-linguistic meanings associated with total, minimal word, heavy syllable, and light syllable reduplication.

11.8 MORPHOLOGICAL LOCATION AND SEMANTIC SCOPE OF REDUPLICATION

Another question one might ask in exploring the derivational character of reduplication is whether reduplication patterns with derivation or with inflection in its affix ordering properties: where in the word does reduplication occur?

As we have seen implicitly throughout this chapter, reduplication can target the entire word, the root, or any stem-sized morphological subconstituent in between; as we have seen, it can even target individual affixes. An explicit illustration of this kind of variation within a language family can be found in Bantu, in which verb reduplication is widespread. The schema in (20), based on work by Downing (e.g. 1997, 1999a, b, 2000, 2006), Hyman (e.g. 2009), and others, shows an internal analysis of the verb which has been motivated in many Bantu languages:



In a study of the natural history of Bantu reduplication, Hyman (2009) identifies examples of reduplication at each level. The semantics of the constructions Hyman surveys are similar, indicating a common historical source. Ciyao (Ngunga 2001) manifests full Stem reduplication, including derivational suffixes (21a) and the final inflectional suffix (21b). By contrast, Ndebele (Sibanda 2004) reduplicates only the Dstem ('derivational stem'), excluding any suffix in the obligatory inflectional FV ('final vowel') position (21c–d). In Kinyarwanda (Kimenyi 2002), only the root is reduplicable, as shown in (21e–f). Verb stems are shown, in all examples in (21), without inflectional or infinitival prefixes, as these do not undergo reduplication:

(21)	Full stem reduplication (all suffixes)	[Ciyao]
a.	telec-el-a 'cook-APPL-FV'	→ telec-el-a + telec-el-a 'cook for someone frequently'
b.	dim-ile 'cultivate-PERF'	→ dim-ile + dim-ile 'cultivated many times'
	Dstem reduplication (no inflectional suffixes)	[Ndebele]
c.	lim-el-a 'cultivate-APPL-FV'	→ lim-e + lim-el-a 'cultivate for/at a little, here and there'
d.	lim-e 'cultivate-SUBJ'	→ lim-a + lim-e (*lim-e + lim-e) 'cultivate a little, here and there (subjunctive)'

Root reduplication (no suffixes)		[Kinyarwanda]
e. rim-w-a	→	rim-aa + rim-w-a (*rim-w-a + rim-w-a)
‘cultivate-PASS-FV’		‘be cultivated several times’
f. rim-ir-a	→	rim-aa + rim-ir-a (*rim-i + rim-ir-a)
‘cultivate-APPL-FV’		‘cultivate for/at, here and there’

Importantly for the question of where reduplication occurs within words, this attenuating or distributive reduplication process occurs inside of most of the productive inflection in the Bantu verb; in the case of root reduplication, it even occurs inside of all of the verbal (valence-changing) derivation. This example illustrates a problem for examining the relationship between reduplication and affix ordering: reduplication very often has wide semantic scope. In its semantics it often patterns with functions that surveys of verbal affix ordering, such as Bybee (1985) or Rice (2000), associate with outer, not inner affixes. Yet reduplication very often targets roots or other internal constituents of words.

A more complicated type of case is presented by languages like Samala (known in the literature as Ineseño Chumash), in which a CVC prefixing reduplication construction which expresses “repetitive, distributive, intensive, or continuative” (Applegate 1972: 383–4) is slotted somewhere within a complex verb whose affixes are descriptively divided into the following zones:

(22) Outer prefixes—Personal prefixes—Inner prefixes—[root—suffixes]_{stem}

Outer prefixes mark things like negative, tense, nominalization/relativization, clause subordination, and sentential adverbs. Personal prefixes are purely inflectional, marking person and number of subject. Inner prefixes are largely derivational, marking a variety of information including aspect, instrumentals, action classifiers, spatial orientation, and verbal force (see Applegate 1972: 301 ff).

As is not surprising given its aspectual meaning, the meaning of reduplication generally scopes over the entire verb, and thus one might expect the CVC reduplicative prefix to occur near the beginning of the word. Instead, reduplication tends phonologically to target the root, as in examples like (23) (Applegate 1972: 387, 1976: 282):

- (23) k-ni-čeq ‘1SUBJ-TRANS-tear’ > kni-čeq-čeq ‘I’m tearing it up’
k-wi-čeq ‘1SUBJ-BY_HITTING-tear’ > kwi-čeq-čeq ‘I pound it to pieces’

Aronoff (1988), Inkelas and Zoll (2005), and others have characterized this process as infixing—an ‘outer’ process, consistent with taking wide semantic scope and being inflectional—whose form appears inside derivational affixes because it is an infix that targets the root. Infixation to the root seems to be especially common among reduplicative affixes (“internal reduplication”). This is not, by contrast, a common pattern for

segmentally fixed affixes, which, when they infix, tend to occupy positions either near the margin of a word or adjacent to a stressed syllable (see e.g. Yu 2007a), not adjacent to a particular morpheme boundary.

The example in (24), from Tagalog, is a particularly clear illustration of the ordering flexibility that inflectional reduplication can have. This particular CVV reduplicative prefix in Tagalog encodes contemplated aspect. It can occur at virtually any location within the string of derivational prefixes, with no effect on meaning. The example in (24) is taken from Rackowski (1999: 5); the general phenomenon of variable reduplicant position in Tagalog is also discussed by Carrier (1979), Condoravdi and Kiparsky (1998) and Ryan (2010), among others:

(24) Unreduplicated		...with contemplated aspect reduplication
ma-ka-pag-pa-hintay	→	ma-[<u>kaa</u> -ka-pag-pa-hintay]
ABILITY-COMplete-TRANS-CAUS		ma-ka- <u>paa</u> -[pag-pa-hintay]
E-WAIT		
'be able to cause someone to wait'		ma-ka-pa- <u>paa</u> -[pag-hintay]
ma-ka-pag-pa- <u>hii</u> -[hintay]		

One possible explanation for the distinctive order properties on the part of partial reduplication may lie in an observation made by Hyman (2009), namely that, possibly for processing reasons, reduplication tends to target root material rather than affixal material. In a number of languages, reduplication occurs on the opposite side of the root from most of the affixes that are in the scope of reduplication. In Bantu languages, verb stem morphology is exclusively suffixing, while verb stem partial reduplication is prefixing. If reduplication is an inflectional prefix in a language with a lot of prefixing derivational morphology, like Samala, the only way to target root material consistently is to be an infix.

The “out-of-place” ordering of reduplication is also facilitated by its tendency to take wide semantic scope. A particularly compelling example of wide-scope reduplication of an inner constituent comes from Harley and Leyva (2009), who discuss internal root reduplication in Hiaki (also known as Yaqui; Uto-Aztecan, Cahita). Habitual reduplication in Hiaki appears to reach into N-V compounds to target the head V but semantically takes scope over the entire compound. Thus the verb *kuta-siute* ‘stick-split = wood-splitting’ reduplicates as *kuta-siu-siute* ‘wood-splitting habitually’; *pan-hooa* ‘bread-make = making bread’ reduplicates as *pan-ho-hoa*; etc. Haugen (2009), like Aronoff (1988) before him, relates head reduplication to the phenomenon of head inflection, familiar from such English examples as *understand* ~ *understood* or *grand-child* ~ *grandchildren*.

We began this section by asking whether the function of reduplication is related to its ordering properties relative to derivation and inflection. Although this question can only be answered on the basis of a broad, genetically and areally balanced

cross-linguistic survey that has not yet been conducted, I propose two generalizations which future research can test. One is that reduplication that has clearly derivational functions, for example changing part of speech, will fairly unambiguously operate on constituents that contain roots and, potentially, other derivational affixes; it will occur inside of inflection. The other is that reduplication whose function falls partially or squarely in the category of inflection is much less constrained in its ordering properties. This is clearly related to the fact that (inflectional) reduplication has wide scope over the whole word, regardless of what part of the word it copies. It is interesting to note a possible connection to morphological negation, which also typically takes wide scope and whose ordering properties are similarly hard to pin down cross-linguistically. More research into these topics is sorely needed.

11.9 CONCLUSION

Although the study of reduplication in the literature has focused particularly on its phonological form and on the question of semantic iconicity, the place of reduplication in a morphological grammar is equally interesting. Reduplication sometimes acts as a “wild card” in morphology, exhibiting combinatoric (affix ordering) behaviors which are uncharacteristic of other morphological constructions. This may be due to the way in which the characteristic iconic semantics of reduplication straddle the boundary between derivation and inflection. Like inflectional morphology, reduplication tends to have wide semantic scope. Like derivational morphology, reduplication tends to alter event-internal meaning. And like derivational morphology, reduplication has a predilection for occurring in phonological proximity to the root. These conflicting factors conspire to paint a fascinating picture.

CHAPTER 12

NON-CONCATENATIVE DERIVATION

Other Processes

STUART DAVIS AND NATSUKO TSUJIMURA

12.1 INTRODUCTION

THIS chapter provides an overview of a wide range of non-concatenative (non-reduplicative) phenomena in morphology focusing on a typological categorization.¹ The definition of non-concatenative morphology is not uncontroversial. Kurisu (2001: 2) considers non-concatenative morphology to be observed in cases where the phonological instantiation of a morpheme cannot be demarcated in an output representation. Bye and Svenonius (2012) similarly define non-concatenative patterns negatively as phenomena that fall short of the concatenative ideal. Briefly, the concatenative ideal entails that the morpheme is segmental (i.e. consists of one or more phonemes), additive (i.e. adds phonological substance to the base), linearly ordered, and contiguous (e.g. prefixes and suffixes). From this perspective, the major phenomena that would be considered non-concatenative are autosegmental affixation (i.e. a morphological category being marked by the addition of a distinctive feature or tone to a base form), infixation, subtractive morphology, and template satisfaction under the view that a morphological template is a segmentally underspecified prosodic node. We frame our overview of non-concatenative morphology in terms of the expression of exponence,² taking

¹ In our presentation of data from a wide variety of languages, we generally preserve the original transcriptions in the sources cited. Transcriptions of data from Standard Arabic and Japanese are based on our own knowledge of these languages and are consistent with what is found in the existing literature. English and German data are presented in their orthographic forms rather than in transcription.

² By exponence, we mean the phonological realization of a morpheme; and for the purpose of our discussion, we specifically focus on exponence of morphemes that are either derivational or inflectional.

non-concatenative morphology to entail morphological processes where exponence is not (exclusively) expressed by the concatenation of additive phonemic content to a base form. This not only excludes the clearly concatenative processes of prefixation and suffixation, but also infixation. Like prefixation and suffixation, infixation processes display consistent phonemic content and infixes can usually be clearly demarcated, thus differing from the non-concatenative phenomena that will be discussed in this overview. We will only touch on infixation when it co-occurs with a non-concatenative phenomenon (e.g. templatic morphology). For a detailed discussion of infixation, see Chapter 9.

In considering a typological categorization of non-concatenative morphology, we make a basic division between two (usually) distinct types: templatic and a-templatic. Templatic morphology involves cases where there are morphological restrictions on the shape of words. In the type of templatic morphology found most commonly in the Semitic languages (e.g. McCarthy 1981, Doron 2003, Bat-El 2011), morphological exponence of a category is expressed by an invariant prosodic shape. A second type of templatic morphology found at least marginally in many languages is instantiated when a concatenative affix imposes a templatic subcategorization requirement on the base to which it attaches. In Section 12.2 we present a variety of examples of templatic morphology that distinguish between these two types. Section 12.3 examines a-templatic non-concatenative morphological processes, and outlines a range of phenomena that include subtractive morphology and moraic augmentation. We also discuss autosegmental affixation in which a distinctive feature is utilized to express exponence as in consonant mutation or vowel change such as umlaut, or in tonal morphology where exponence is expressed by a certain tone or tone pattern. Section 12.4 concludes the chapter by briefly considering some of the theoretical issues related to non-concatenative morphology. While the focus of this volume is primarily dedicated to the theme of derivational morphology, the question arises as to whether non-concatenative morphology can be subsumed under derivational or inflectional morphology since the phenomena pertinent to non-concatenative morphology in this chapter have resemblance to derivational morphology in some cases and to inflectional morphology in others. We will touch on this issue at the end of the chapter. (See Chapter 2 of this volume for a discussion of the problems of distinguishing between the two.)

Before we begin, it is important to ask whether non-concatenative derivation is theoretically distinct from concatenative morphology in such a way that it requires a different formal mechanism, or whether the distinction is epiphenomenal, as Bye and Svenonius (2012) maintain. The contrastive views can be seen as the modern incarnation of Hockett's (1954) distinction between item-and-arrangement vs. item-and-process morphology. Bermúdez-Otero (2012) refers to current theories that view morphology from an item-and-arrangement perspective as "piece-based" theories and those that take a processual view as "process-based" theories. Non-concatenative morphology appears more compatible with process-based approaches whereas concatenative morphology

Typically this is achieved by adding consistent phonemic content to a base in order to express the semantic content of a given morpheme.

is more in line with piece-based theories. Some recent perspectives on morphology, however, attempt to unify concatenative and non-concatenative morphology under a single theoretical approach. From the strictly pieced-based view, Bye and Svenonius (2012) argue that non-concatenative morphology is theoretically epiphenomenal: all affixation is contentful, but non-concatenative effects can arise because affixes may be deficient featurally or even segmentally (as maintained in work like Lieber 1984, 1987); on the other hand, it can arise if the relation of an affix to higher level prosodic structure is pre-specified (resulting in templatic effects). Under the process-based theory of Anderson's (1992) a-morphous morphology, morphology is a process acting on stems or words to produce complex forms. According to this theory, processes reflect those found in phonology such as deletion (e.g. subtractive morphology), featural change (e.g. mutation, umlaut), and lengthening (e.g. moraic augmentation). For example, the perfective in the Uto-Aztecan language Tohono O'odham is typically formed from the imperfective base by the deletion of the final consonant. This would be formally expressed by the rule of perfective formation, which would delete the final consonant of the base. Common prefixation and suffixation phenomena reflect rules that introduce the phonemic exponence of an affix (i.e. the phonemic sequence that comprises an affix) as part of the rule for the morphological process. For instance, the regular English plural rule would introduce /-z/ to a noun base. Here, too, we see that non-concatenative morphology can also be viewed as theoretically epiphenomenal: it arises as the result of the type of rule that the morphological process requires.³ There is no formal distinction between the non-concatenative subtractive morphology of the Tohono O'odham perfective and the concatenative English plural other than that they entail different rules. Even within Optimality Theory, there is a division between the piece-based and process-based approaches to non-concatenative morphology. The former is shown in Wolf (2007), Bye and Svenonius (2012), and the stratal OT approach of Bermúdez-Otero (2012, forthcoming), while the latter is developed in the anti-faithfulness theory of Alderete (1999, 2001) and in the morpheme realization theory of Kurisu (2001). The purpose of this chapter is not to resolve these controversies, but instead to overview the range of phenomena as instantiations of non-concatenative morphology that give rise to the controversy.

12.2 TEMPLATIC MORPHOLOGY

Since the seminal work of Chomsky and Halle (1968) (henceforth *SPE*), two avenues of phonologically based research have played an important role in the emergence

³ It should be noted that not all current theories of morphology are strictly "piece-based" or "process-based." For example, construction morphology (Booij 2010) can allow for both "pieces" and "processes" as part of a morphological construction. As noted by Tsujimura and Davis (2011a, b), a prosodic template can be a basic part of a form-meaning pairing of a constructional schema in construction morphology.

of the study of non-concatenative processes: autosegmental phonology (Goldsmith 1976) and prosodic morphology (McCarthy 1984, McCarthy and Prince 1986). Autosegmental phonology offered a formal means to analyze morphological processes in which the exponence is partially subsegmental as in mutation or umlaut (Lieber 1984, 1987, 1992). Prosodic morphology introduced a way to deal with morphological processes that are characterized by invariant templatic shape. While languages with non-concatenative morphology have long been known, the theoretical constructs of autosegmental phonology and prosodic morphology allowed researchers in the 1980s and 1990s to focus on non-concatenative processes, leading to new formal approaches. Work on templatic morphology in the post-*SPE* period originates in McCarthy's (1979, 1981, 1984) seminal research on Arabic and Hebrew with extended development of a prosodic theory of templatic morphology in McCarthy and Prince (1986, 1990). McCarthy's original work offered an autosegmental analysis of Semitic root-and-pattern morphology in which the CV pattern of a word could constitute a separate morpheme since the shape of the word is a crucial component to the meaning. Subsequent work by McCarthy and Prince (1986, 1990) introduced a constrained theory of templatic shapes in morphology, which maintained that morphological templates always constitute authentic units of prosody such as a syllable or a foot. This became known as the "Prosodic Morphology Hypothesis," and it gave rise to a research program that was especially active in the 1980s and 1990s and continues today, though under different guise.

While McCarthy's early work focused on the root and pattern morphology of Semitic in which the prosodic template itself contributes to the meaning of the word, Archangeli (1983) identified a somewhat different type of templatic morphology, one where an affix imposed a particular templatic shape on the base to which it attached. She showed that some suffixes in the Penutian language Yawelmani required their verb base to have a certain prosodic shape (e.g. CVCC and CVCVVC). The shape of the verb base then could change depending on the suffixal imposition. Such cases differ from the Semitic type in that exponence in Yawelmani can be viewed as purely concatenative with suffixes subcategorizing for a certain prosodic shape. That is, the prosodic shape does not uniquely instantiate morphological exponence. This latter type of templatic morphology is found at least marginally in many languages. Section 12.2.1 presents examples of the Semitic type of templatic morphology where exponence of a category is expressed solely by an invariant prosodic shape. Section 12.2.2 presents cases where a concatenative affix imposes a templatic requirement on the base to which it attaches.

12.2.1 Template as Morpheme Exponence

Templatic morphology is pervasive in both nouns and verbs of Semitic languages, but the verbal systems are particularly striking because of their root-and-pattern system of non-concatenative morphology. Basic verb forms are not comprised of contiguous

morphemes but show interleaving of elements.⁴ Consider the three forms in (1) related to the meaning of “write” found in Standard Arabic and the three forms in (2) related to the meaning “kill.”

(1) Arabic verb forms—“write”

	<i>Verb</i>	<i>Template</i>	<i>Gloss</i>	<i>Passive</i>
a.	katab	CVCVC	‘wrote’	kutib
b.	kattab	CVCCVC	‘dictated (write, causative)’	kuttib
c.	kaatab	CVVCVC	‘corresponded (write, reciprocal)’	kuutib

(2) Arabic verb forms—“kill”

	<i>Verb</i>	<i>Template</i>	<i>Gloss</i>	<i>Passive</i>
a.	qatal	CVCVC	‘kill’	qutil
b.	qattal	CVCCVC	‘massacre (intensive)’	quttil
c.	qaatal	CVVCVC	‘battle one another (reciprocal)’	quutil

A morphological analysis of (1) and (2) would show that the consonants *ktb* and *qtl* provide the lexical meaning “write” and “kill,” respectively. The vowel pattern provides grammatical information (*a* and *ui* indicate past tense and passive, respectively), and the overall word shape seems to add meaning such that CVCCVC corresponds to causative and/or intensive (cf. Doron 2003) and CVVCVC marks reciprocal. In an analysis like that of McCarthy (1979, 1981), vowels and consonants are represented separately on different morphological tiers since both the vowel pattern and consonant sequence comprise morphological entities. Furthermore, since the specific CV pattern of the verb contributes to its meaning, a specific CV template encoding the shape is represented on a separate tier to which the consonants and the vowels are linked. To illustrate the analysis, *kaatab* ‘corresponded with’ in (1c) is formed on the basis of the reciprocal template, CVVCVC, with the consonantal tier consisting of *k-t-b* to give the meaning of “write” and the vowel tier *a*, which is the tense/aspect/mood marker. The different morphological tiers are shown in (3).

(3) Base form: katab “write”

k	t	b	← "write"
C	V	V	C
	\	\	/
	a		← past, active
→ kaatab "corresponded with"			

⁴ Person marking on Semitic verbs is expressed through concatenative affixation. As is typical in discussion on Semitic verbal templates, sample forms that we present are from the unmarked form, i.e. the masculine third person past tense.

Semitic morphology is unusual in that the consonantal sequence and vowel pattern can be analyzed as comprising separate morphemes even though they are intertwined and not kept separate in the actual pronunciation. As noted by Bat-El (2011), this type of root-and-pattern morphology is pervasive in Semitic languages. Emerging from McCarthy's (1979, 1981) templatic analysis of Semitic was the issue of possible templatic shapes. If the morphological template is expressed in terms of CV-slots as in (3), then hypothetically there is no restriction on a template shape. It could be any combination of C-slots and V-slots. Beginning with McCarthy (1984), there was a reanalysis of the CV template as referring to higher level prosodic constituents such as syllables and feet. This is illustrated by the Modern Hebrew form II verbal construction, referred to as the *piel* in traditional Hebrew grammar. Verbs of this class are normally causative or intensive, but are also denominal or neologisms (frequently from borrowings). A full discussion of its semantics is found in Doron (2003), who labels verbs of this class as having the intensive template, although she shows that not all verbs of this class have intensive meaning (i.e. there are other reasons for denominals and neologisms to fall into this class). Formally, this class of verbs in Hebrew is distinct in that they typically have the vowel sequence /i/-/e/ to indicate that the verb is past tense (active). In the data in (4), five subclasses of the form II verb can be observed based on their CV pattern. We also indicate the base of the form II verb, which can be either a simple verb or a noun.

(4) Modern Hebrew form II subclasses

Subclass 1: CVCVC

a.	limed	'teach'	base: lamad 'learn'
b.	tiken	'repair'	base: takan 'be straight'
c.	rikez	'concentrate'	base: merkaz 'center'
d.	yiven	'Hellenize'	base: yavan 'Greece'

Subclass 2: CVCCVC

a.	tirgem	'translate'	base: targum 'translation'
b.	kifter	'button'	base: kaftor 'button'
c.	tilpen	'telephone'	base: telefon 'telephone'
d.	ʔixzev	'disappoint'	base: ʔaxzava 'disappointment'

Subclass 3: CVCCCVC

a.	tilgref	'telegraph'	base: telegraf 'telegraph'
b.	sinkren	'synchronize'	base: sinkroni 'synchronic'
c.	tirklen	'arrange a room'	base: t(e)raklin 'room'
d.	sindler	'cobble'	base: sandlar 'cobbler'

Subclass 4: CCVCCVC

a.	flirtet	'flirt'	base: flirt 'flirt'

Subclass 5: CCVCCCVC

a.	stingref	'take shorthand'	base: stenografit 'stenographer'

From a perspective of CV sequences, there are five different templatic shapes in (4), but by reference to units of prosody, all five subclasses can be collapsed into a single prosodic template consisting of two syllables or a foot. We can then schematize the Form II pattern with the example in (5), where each tier represents a different morpheme.⁵

(5) Base: kaftor “button”

k	f	t	r	button
\ /		\ /		
σ		σ		prosodic template (Form II, intensive, denominal, neologism)
i		e		past, active
→ kifter				

Consequently, we see from the Hebrew form II example that reference to higher prosodic structure like syllable and foot can unify apparently different CV shapes into a single form.⁶

In addition to verbal examples, many Semitic nominal derivations can also be viewed as having prosodic templates. A pertinent example comes from a common pattern of hypocoristic (nickname) formation in Arabic, as is discussed by Davis and Zawaydeh (1999a, 2001). The hypocoristic adds a sense of endearment since they are normally used among family members or intimates and not in front of outsiders. While the pattern illustrated in (6) is widespread in Arabic dialects, our presentation is based on the Ammani-Jordanian Arabic, discussed by Davis and Zawaydeh (1999a, 2001).

(6)	<i>Full Name</i>	<i>Hypocoristic</i>	<i>Full Name</i>	<i>Hypocoristic</i>
a.	xaalid	xalluud	e.	widaad
b.	basma	bassuum	f.	maryam
c.	saliim	salluum	g.	ibraahiim
d.	bufra	baffuur	h.	muusa
				masmuus

⁵ The details of the mapping that produces [kifter] as opposed to *[kftir] is left for the phonology to determine. Also, as Bat-El (1994) and Ussishkin (1999) have pointed out, there can sometimes be transfer effects from a base noun in denominal examples such as with the denominal of “flirt” as [flirtet] rather than *[fliret]. We do not discuss the tier conflation process where linearization takes place to produce the surface output. This was much debated in the 1980s. See Ussishkin (2011) for some discussion.

⁶ This raises the question regarding the Arabic data in (1) and (2), where the three different verbal classes are all bisyllabic but with distinct first syllables. With CV templates, the three classes are CVCVC, CVCCVC, and CVVCVC. A prosodic template expressed in terms of syllables or feet would have to indicate, in some way, the nature of the first syllable. McCarthy (1993) makes the interesting suggestion that the CVCCVC pattern can be analyzed as having a CVCVC template along with affixation of a (consonantal) mora.

Regardless of the phonological shape of the first name, the hypocoristic always has the same bisyllabic templatic shape in which the first syllable is closed and the second syllable has a long vowel. For convenience, we represent this as $C_iVCCVVC_f$, where C_i is the initial consonant of the full name and C_f is the final consonant of the full name. The vowel of the first syllable of the hypocoristic template is specified as /a/ and that of the second syllable as /u/, which is realized as long. The data in (6a–e) show that in names with three consonants, the medial consonant of the full name is realized as a geminate in the hypocoristic. The data in (6f–g) illustrate that the hypocoristic template can accommodate names that have four consonants, while the name in (6h) indicates that the template can also handle names with only two consonants by consonantal reduplication. Setting aside the phonological issue of how the mapping is realized between the full name and the hypocoristic form, we can exemplify the Ammani-Jordanian Arabic hypocoristic as in (7). (See Davis and Zawaydeh (1999b) for a detailed formal analysis.)

(7) Base name: basma

b	s	m	(root consonants)
	/ \		
C	V	C C V V C	hypocoristic
		\	
a		u	(vocalic melody)

→ bassuum

It is not common to find templatic morphology of the type discussed in this section (i.e. where the prosodic shape of the form contributes to its meaning) outside of Semitic languages (or perhaps Afroasiatic more generally). When it does occur, it is not pervasive in the language but is characteristic of isolated constructions. We will demonstrate this with two well-discussed examples: the Rotuman incomplete phase and the Cupeño habitative. In the Austronesian language Rotuman, the incomplete phase of a word is formed from the complete phase by a variety of processes. This is shown in (8): the data are taken from McCarthy (2000: 148), based on Churchward (1940).

(8) Rotuman phase alternations

<i>Process</i>	<i>Complete</i>	<i>Incomplete</i>	<i>Gloss</i>
a. Deletion	tiʔu	tiʔ	'big'
	sulu	sul	'coconut-spathe'
	rako	rak	'to imitate'
b. Metathesis	iʔa	iaʔ	'fish'
	hosa	hoas	'flower'
	parofita	parfiat	'prophet'
	pure	puer	'to rule'
c. Umlaut	mosi	mös	'to sleep'
	futi	füt	'to pull'

d. Diphthongization (. indicates syllable boundary)

le.le.i	le.lei	'good'
ke.u	keu	'to push'

e. No alternation

rii	rii	'house'
si.kaa	si.kaa	'cigar'

At first glance there may seem to be no obvious exponence of the Rotuman incomplete phase; rather, it is purely processual. This process could be deletion, metathesis, umlaut, diphthongization, or no alternation, depending on the lexical item. However, as McCarthy and Prince (1986) and McCarthy (2000) show, there is a homogeneity in the nature of the output in the Rotuman incomplete phase: the final syllable must be heavy (i.e. bimoraic, ending in a long vowel, diphthong, or final consonant). From a templatic perspective, it can be posited that a bimoraic syllable (or monosyllabic foot) template is the exponence of the incomplete phase. The actual process that the input undergoes in order to meet the templatic requirement is phonologically determined based on the nature of the input vowels. Where vowel sequences or (light) diphthongs are possible, metathesis in (8b) or diphthongization in (8d) will occur. If metathesis or diphthongization cannot occur because the output of such processes would result in an impermissible vowel sequence, then deletion takes place either with concomitant umlaut if the deleted vowel is front, as in (8c), or without umlaut, as in (8a). If the final syllable is already bimoraic as in (8e), there is no distinct form for the incomplete phase. However, such forms have exponence of the incomplete phase (bimoraic syllable), although the mapping of the input to the bimoraic template in (8e) does not result in an output form that is different from the complete phase. This, from a descriptive perspective, constitutes an example outside of Semitic where a prosodic template marks exponence. (See McCarthy (2000) for a detailed analysis of the Rotuman incomplete phase.)

The habilitative in the Uto-Aztecan language Cupeño (cf. Hill 1970, Hill and Nolasquez 1973)—another non-semitic language—provides an intriguing example whereby the exponence of a morphological process is expressed by a specific templatic shape, but at the same time the process is restricted in its application to words that have a certain phonological characteristic (cf. McCarthy and Prince 1986, 1990, Crowhurst 1994). Consider the data in (9), taken from Crowhurst (1994) (mainly based on Hill (1970) and Hill and Nolasquez (1973)).

(9) Cupeño Habilitative

<i>Verb Stem</i>	<i>Habilitative</i>	<i>Gloss</i>
a. čál	čáʔaʔal	'husk'
táw	táʔəʔəw	'see'
həlʔáp	həlʔəʔəʔəp	'hiccup'
b. páčik	páčiʔik	'leach acorns'
čáŋnəw	čáŋnəʔəw	'be angry'
čəkúk ^{wilʔ}	čəkúk ^{wiʔilʔ}	'joke'

c.	pínəʔwəx	pínəʔwəx	'sing enemy songs'
	xáləyəw	xáləyəw	'fall'
d.	čí	číʔ	'gather'
	ʔáyu	ʔáyu	'want'
	ʔiyú:nə	ʔiyú:nə	'fast'

There is a basic distinction between the data in (9a–c) and (9d). The habilitative forms in (9a–c) all end in the same prosodic trisyllabic sequence of a stressed syllable of the verb stem followed by two stressless syllables. McCarthy and Prince (1986, 1990) express this as a trisyllabic foot template headed by the stressed syllable.⁷ That is, from a purely descriptive point of view, the exponence of the habilitative consists of a trisyllabic template. If the verb stem ends in a stressed syllable as in (9a), two epenthetic syllables are added. The forms in (9b) demonstrate that if the verb stem contains a stressed syllable followed by one stressless syllable, then the habilitative contains one epenthetic syllable. Those in (9c) show that if the verb stem already contains a stressed syllable followed by two stressless ones, then the habilitative is identical to the stem itself. As with the unchanged incomplete phase of Rotuman in (8e), there is morphological exponence with the habilitative in (9c): it simply does not have a distinct realization from the verb stem. Given the expression of morphological exponence in (9a–c) through the trisyllabic (foot) template, the data in (9d) are noteworthy. In these words, the habilitative does not end in a trisyllabic sequence as specified by a template; instead, the habilitative form is essentially the same as the verb stem. The key difference between the verb stems in (9d) and those in (9a–c) is that the former ends in a vowel rather than a consonant. This indicates that there is an input restriction on the application of the Cupeño habilitative, namely, it applies only to verb stems that end in a consonant; otherwise, the verb does not have a unique exponence for the habilitative. The application of the trisyllabic habilitative template in Cupeño is thus restricted to verb stems that end in a consonant.⁸

The Rotuman incomplete phase and the Cupeño habilitative have provided two instances of the use of morphological templates outside of the Semitic languages. They are, however, different from the Semitic examples discussed earlier in that the domain of the template in Semitic applies over the entire output word. While this is the pattern observed with most of the Cupeño and Rotuman data given above, examples like the Rotuman word for “prophet” in (8b) as well as the Cupeño word for “hiccup” in (9a) and “joke” in (9b) show that in longer words, the initial syllable may be outside the domain of the morphological template, which nevertheless is expressed. The Cupeño habilitative further demonstrates an instance of an input requirement to a morphological process, that is, that the verb has to end in a consonant for the distinct manifestation of the habilitative template.

⁷ This characterization is different from Crowhurst (1994), who describes it as a bisyllabic template following the stressed syllable.

⁸ The addition of a glottal stop to the CV stem in the first example of (9d) can be viewed as reflecting a minimal bimoraic requirement on surfacing words (cf. Crowhurst 1994: 197).

12.2.2 Template as Prosodic Subcategorization Requirement on Concatenative Affixation

Much of the literature on templatic morphology documents examples where concatenative affixes impose prosodic templatic requirements on the base to which they attach. These cases are different from the Semitic type cases since the morphological exponence is expressed through the surfacing affix, not necessarily the template. The templatic requirement on affixation can be expressed through subcategorization frames associated with specific affixes, as in Booij and Lieber (1993) and Bermúdez-Otero (2012), by prosodic circumscription as in McCarthy and Prince (1990), or as affix-specific alignment constraints in Optimality Theory. It is then interesting to examine the outcome of affixation processes where a prosodic template is imposed upon the base, especially when there is a conflict between the prosodic structure of the base and the templatic shape imposed by the affix. We can discern three different outcomes in such processes in descriptive terms: (i) the base may change its shape to fit the templatic requirement of the affix, (ii) affixation may fail to occur if the base does not have the particular prosodic shape for which the affix subcategorizes, or (iii) the affix is “mobile”, seeking out the particular prosodic shape somewhere within the base, and often results in infixation. We give examples of each of these three types.

The California Penutian language Yawelmani, as discussed by Archangeli (1983, 1985, 1991) and Inkelas (2011), presents a telling example of how a base word can reconfigure its shape in order to match a prosodic templatic requirement that is imposed by a suffix. Many suffixes in Yawelmani impose templatic requirements on the stem. Consider the two verbs in (10a–b) given with their underlying base forms. (The nature of the underlying form is known from suffixes that do not impose a templatic requirement.) The data in (10c–d) show the two verbs with the reflexive/reciprocal suffix /-iwsuul/, while the examples in (10e–f) demonstrate the same verbs with the durative suffix /-iixok/. Surface pronunciations are given in brackets. We do not discuss the phonological changes determined by processes of vowel harmony and long vowel lowering.

(10) Yawelmani template selecting affixes

a. /luk'l/ 'bury'	b. /huluus/ 'sit'
c. /luk'l-iwsuul/ → [luk'ool-uwsool]	d. /huluus-uwsool/ → [huloos-uwsool]
e. /luk'l-iixok/ → [lik'l-eexok]	f. /huluus-iixok/ → [huls-eexok]

The reflexive/reciprocal suffix in (10c–d) requires that its base have an iambic shape, CVCVVC. Its effect is seen clearly in (10c) where the underlying CVCC verb form acquires an extra vowel that is lengthened in order to match the prosodic requirement of the suffix. No (prosodic) change is seen in (10d) since the verb is underlyingly iambic. On the other hand, the suffix /-iixok/ in (10e–f) requires that its verbal base have the templatic shape CVCC. Here, the change is witnessed in (10f), where the underlying

verb /huluus/ loses its long vowel in order to satisfy the template. It is clear from the specific examples shown in (10) that the imposition of the prosodic template on a base is unrelated to the phonological structure of the affix: both affixes are bisyllabic with a heavy first syllable. Yawelmani constitutes an interesting example in which the underlying configuration of base words can be altered because of a prosodic requirement that a suffix imposes.

A somewhat different case of a suffix imposing a template is the Japanese hypocoristic suffix *-tyan*. Although the suffix can attach to any full first name as seen in (11), interesting patterns emerge when it attaches to a truncated name as seen in (12). Data are from Poser (1990) and Tsujimura (2007) as well as the intuitions of the second author.

(11) Japanese hypocoristic *-tyan* with full names

- a. satiko → satiko-tyan b. akiko → akiko-tyan
c. masao → masao-tyan d. syuusuke → syuusuke-tyan

(12) Truncated Japanese hypocoristics with *-tyan*

- a. satiko: sat-tyan, saa-tyan, sako-tyan (*sa-tyan, *satit-tan)
b. akiko: at-tyan, aa-tyan, aki-tyan, ako-tyan (*a-tyan, *akit-tyan)
c. masao: mat-tyan, maa-tyan, masa-tyan (*ma-tyan, *masat-tyan)
d. syuusuke: syuu-tyan (*syu-tyan, *syuusu-tyan)

In truncated hypocoristics with the suffix *-tyan*, the base name must have exactly two moras. That is, the suffix imposes a bimoraic foot upon the base name. The truncated hypocoristics in (12) demonstrate that the templatic requirement can be achieved in a variety of ways, but that one-mora and three-mora hypocoristic forms do not occur, as is expected by the bimoraic requirement. The Japanese example is similar to the Yawelmani case in that a suffix that expresses the morphological exponence imposes a templatic requirement on the base. It is different from Yawelmani, however, because there is optionality in how the template is satisfied. Nevertheless, the suffix itself occurs outside of the template in both cases.

It appears that suffixal hypocoristic formation in a variety of languages can impose a template on the base that has the effect of shortening it. This tendency applies to English. The common *y*-hypocoristic is formed by suffixing [-i] (often spelled as “y”) to a syllable. Consider the data in (13).

(13) English *y*-hypocoristics

- a. Susan → Susie d. Patricia → Patty g. Martin → Marty
b. Timothy → Timmy e. Gabriella → Gabby h. Barbara → Barbie
c. Kenneth → Kenny f. Ignatius → Iggy i. Sandra → Sandy

The hypocoristic suffix is added to the initial syllable of the base name, but the syllable to which the hypocoristic suffix *y* attaches does not necessarily correspond to the syllable

as it appears in the full name. In *Patricia* in (13d), the /t/ is at the beginning of the second syllable, but it is included in the base name. If one just considers the sequence of phonemes that make up the full name *Patricia*, the maximal initial syllable is the sequence *pat*; this is the form to which the hypocoristic -y is affixed in order to derive *Patty*. Similarly with the name *Sandra* in (13i), the maximal initial syllable given the sequence in the full name is *sand*. The hypocoristic suffix -y follows this sequence to derive *Sandy* even though the /d/ in the full name *Sandra* belongs to the second syllable. Thus, we see that the hypocoristic suffix -y imposes a syllable template on its base that is independent of the syllabification of the full name. We view the English y-hypocoristic formation as involving a suffix that imposes a certain prosodic shape on the base where the syllable prosody of the base can be reconfigured to satisfy the template.⁹

Unlike Yawelmani verb suffixation and hypocoristic formation in Japanese and English, there are cases in which the base does not change its form in order to fulfill a templatic requirement of an affix. For instance, the suffix -er/-est in English comparative/superlative constructions requires that the base adjective be no more than two syllables (with individual variation on the acceptability of some two syllable forms as noted by Carstairs-McCarthy (1998)). Representative data are given in (14):

(14)	<i>Adjective</i>	<i>Comparative</i>	<i>Adjective</i>	<i>Comparative</i>	
a.	smart	smarter	e. intelligent	*intelligenter	(more intelligent)
b.	funny	funnier	f. hilarious	*hilariouser	(more hilarious)
c.	simple	simpler	g. elementary	*elementrier	(more elementary)
d.	pretty	prettier	h. beautiful	*beautifuler	(more beautiful)

The forms in (14e–h) show that if the base does not conform to the two-syllable requirement, the comparative form with -er is impossible. This is quite different from the English y-hypocoristic case where y is suffixed to the initial syllable of the base and all subsequent base phonemes are not realized. If the English -er comparative were like the hypocoristic formation, adjectives of greater than two syllables could participate in the construction with a shortened base form. For example, *intelligent* would truncate to

⁹ A full analysis of the English hypocoristic pattern can be found in Lappe (2007), which includes technical questions on the form of English y-hypocoristics involving full names with obstruent clusters (e.g. *Victoria/Vicky*, *Christina/Chrissy/Christy*). Lappe convincingly argues that the base of the English y-hypocoristic is the full name and not the truncated version of the name. For example, *Timothy* not *Tim* serves as the base of the hypocoristic *Timmy*. Lappe (2007) explains that there are quite a few instances of mismatch between the truncated name and its y-hypocoristic counterpart. For example, the truncated name of *Susan* is *Sue* but the y-hypocoristic is *Susie*, not **Suey*. Similarly, the y-hypocoristic of *Sandra* is *Sandy* although there is no truncated name *Sand* for *Sandra*. With respect to English

intell to form **inteller*; but this does not occur. Thus, the nature of templatic satisfaction varies depending on the morphological process.

In the literature on prosodic morphology (e.g. McCarthy and Prince 1990), there is yet another type of case where a morphological operation applies to a prosodically delimited part of a base word. One well-cited example concerns the possessive affix *-ka* in the language Ulwa spoken in Nicaragua. The data in (15) are cited by McCarthy and Prince (1990).

(15) Ulwa possessive

<i>Base noun</i>	<i>Possessed form</i>	<i>Gloss</i>
a. sana	sana-ka	'deer'
b. bas	bas-ka	'hair'
c. kii	kii-ka	'stone'
d. al	al-ka	'man'
e. amak	amak-ka	'bee'
f. sapaa	sapaa-ka	'forehead'
g. suulu	suu-ka-lu	'dog'
h. kuhbil	kuh-ka-bil	'knife'
i. baskarna	bas-ka-karna	'comb'
j. siwanak	siwa-ka-nak	'root'
k. karasmak	karas-ka-mak	'knee'
l. anaalaaka	anaa-ka-laaka	'chin'

Attention should be paid to what determines the location of the possessive affix *-ka*. While (15a–f) suggest it is a suffix, the data in (15g–l) instantiate infixation: *-ka* occurs as an infix either after the first syllable (15g–i) or after the second syllable (15j–l). McCarthy and Prince (1990) observe that these different environments can be unified if it is posited that *-ka* is suffixed to an initial iambic foot. Such a foot would consist either of an initial heavy syllable as in (15b–d, g–i) or of the initial bisyllabic sequence as exemplified in (15a, j–l). Alternatively, it can be considered that Ulwa possessive *-ka* suffixation is similar to the English comparative *-er* suffixation in that both impose a prosodic template calculated from the left edge of the base word. For the English comparative, if the base does not exactly meet the template, *-er* comparative fails to apply. In the Ulwa case, as seen in (15g–l), the suffix moves to after the initial foot, surfacing as an infix. If the English *er*-comparative were to incorporate the same strategy as Ulwa, the comparative form of *intelligent* would be expected to surface as **intell-er-igent*, contrary to fact.

truncated names without the *y*-hypocoristic suffix and English truncated words more generally (e.g. *fridge* for *refrigerator*), there is an ongoing debate as to whether they reflect morphological operations or are extra-grammatical (i.e. not part of the grammar). We take the view of Dressler and Merlini Barberesi (1994) that English truncated words without an affix are outside of the grammar. See Alber and Arndt-Lappe (2012) for extensive discussion of the different perspectives.

Consequently, we see from our survey that there are a variety of outcomes when a suffix imposes a prosodic requirement on a base that is in conflict with the actual prosody of the base form.

We have thus far classified templatic morphology into two types: (i) the templatic shape expresses exponence, and (ii) the template can be viewed as a prosodic requirement on affixation. Nothing precludes both types of templatic morphology from occurring in the same process and, although not common, the Arabic broken plural (McCarthy and Prince 1990) and the Choctaw *y*-grade (Lombardi and McCarthy 1991) serve to illustrate the co-occurrence of the two types. We briefly discuss the former. The Arabic broken plural pattern is illustrated in (16), taken from McCarthy and Prince (1990). (Also see Spencer 1991.)

(16) Arabic broken plural

	<i>Singular</i>	<i>Plural</i>	<i>Gloss</i>
a.	nafs	nufuus	'soul'
b.	rajul	rijaal	'man'
c.	xaatim	xawaatim	'signet-ring'
d.	jundub	janaadib	'locust'
e.	taqdiir	taqaadiir	'calculation'

While the plural patterns of the nouns in (16) look diffuse, McCarthy and Prince (1990) point out that they all begin with an initial iambic sequence (i.e. CVCVV), suggesting that an iambic template expresses the exponence of the plural. Furthermore, it is noteworthy as to which consonantal phonemes are realized in the iambic template (CVCVV) of the plural. (We ignore the issue of the vowel pattern of the plural.) Compare, for example, (16b) and (16d) with (16c). In both (16b) and (16d) the second consonant of the base noun is realized as the second consonant of the CVCVV iambic template of the plural, whereas the second consonant of the base noun *xaatim* in (16c) is not realized as the second consonant of the corresponding plural, *xawaatim*. In McCarthy and Prince's (1990) analysis, the initial trochaic sequence of the base noun maps onto the iambic template. In (16b) and (16d), where the initial vowel of the singular is short, the second consonant is contained within the initial trochaic sequence. In contrast, the vowel of the first syllable of the singular base in (16c) is long and the second consonant of the singular base is outside of the initial trochaic sequence, so it cannot map onto the iambic template. (The [w] that occurs at the beginning of the second syllable of the plural in (16c) can be viewed as a kind of default consonant.) In this way, we can analyze the Arabic broken plural as having exponence expressed templatically (as an iambic foot) and also subcategorizing for a templatic sequence (a trochaic foot), which maps onto the iambic foot. Such complex cases of templatic morphology suggest that morphemes displaying templatic exponence can combine with a templatic subcategorization requirement. (See McCarthy and Prince (1990) for a detailed analysis of the Arabic broken plural.)

12.3 A-TEMPLATIC NON-CONCATENATIVE MORPHOLOGY

A-templatic non-concatenative phenomena that we will discuss in this section refer to the situation in which morphemic exponence may not have any consistent phonemic realization, but is either subtractive, augmentative, or involving autosegmental affixation. These seem more compatible with processual theories of morphology (“morphology-as-process”) as in Anderson’s (1992) A-morphous Morphology or Kurisu’s (2001) Realize Morpheme theory, which is couched within Optimality Theory. Contrastive to these, some current work, like that of Trommer and Zimmermann (2010), Zimmermann and Trommer (2011), and Bye and Svenonius (2012) approach the phenomena that fall under the purview of a-templatic non-concatenative morphology by maintaining a “morpheme-as-pieces” view. We will focus on the most important phenomena for which different approaches need to account, without detailing specific arguments pertinent to each position. Included in our survey are subtractive morphology, moraic augmentation, and various instantiations of autosegmental morphology where exponence can be realized by a change in a distinctive feature value (e.g. consonant mutation, umlaut) or by tonal imposition.

12.3.1 Subtractive Morphology

Subtractive morphology occurs when a morphological class is marked by deleting a phoneme or some phonemic sequence from the base. Clear illustrations are found in three native American languages: Tohono O’Odham (also known as Papago, Uto-Aztecan), Koasati (Muskogean), and Alabama (Muskogean). The subtraction processes in these languages are similar but not identical. First, consider the Tohono O’Odham perfective forms in (17–18), which show that the perfective is derived from the imperfective in one of three ways. Data are cited from Anderson (1992), Yu (2000), and Horwood (2001), based on Zepeda (1983) and Hill and Zepeda (1992).

(17) Tohono O’Odham perfective—deletion

Subclass 1—final consonant deletion

	<i>Imperfective</i>	<i>Perfective</i>	<i>Gloss</i>
a.	ñeok	ñeo	‘spoke’
b.	bisck	bisc	‘sneezed’
c.	ma:k	ma:	‘gave’
d.	him	hi:	‘walked’
e.	sikon	siko	‘hoed’
f.	hi:nk	hi:n	‘barked’

Subclass 2—final rhyme (VC) deletion

	<i>Imperfective</i>	<i>Perfective</i>	<i>Gloss</i>
g.	ceposid	cepos	'branded'
h.	keliw	kel	'shelled corn'
i.	bijim	bij	'turned around'
j.	huDuñ	huD	'descended' (note: D=retroflex)

(18) Tohono O'Odham perfective, Subclass 3—no deletion

	<i>Imperfective</i>	<i>Perfective</i>	<i>Gloss</i>
a.	gagswua	gagswua	'combing'
b.	dada	dada	'arriving'
c.	mu:	mu:	'wounding by shooting'
d.	bia	bia	'dishing out food'

In Subclass 1, the perfective is formed from the imperfective by deletion of a final consonant. In Subclass 2, the final consonant also deletes; in addition, there is a phonological process that deletes a final high vowel when it is after a coronal (Hill and Zepeda 1992). The Subclass 3 forms in (18) illustrate that when there is no final consonant, the word does not have a distinct perfective form. In this language the exponence for the perfective seems to be the deletion process itself. Notice that no template is necessary because the perfective form can be of any length and the final syllable can either be light (as in (17e)) or heavy.

A slightly different pattern of subtractive morphology is found in the plural verb formation of Koasati, as demonstrated in (19). Data are taken from Horwood (2001) based on Martin (1988) and Kimball (1991).

(19) Koasati—plural verb formation

	<i>Singular</i>	<i>Plural</i>	<i>Gloss</i>
a.	latáf-ka-n	lát-kan	'to kick something'
b.	misíp-li-n	mís-li-n	'to blink'
c.	iyyakkohóp-ka-n	iyyakkóh-ka-n	'to trip'
d.	tipás-li-n	típ-li-n	'to pick something off'
e.	icoktaká:-li-n	icokták-lin	'to open one's mouth'
f.	acití:-li-n	acít-li-n	'to tie something'
g.	facó:-ka-n	fás-ka-n	'to sleep with someone'

The plural is formed from the singular by deleting the rhyme of the penultimate syllable. The deleted rhyme comprises a VC sequence in (19a–d) and a long vowel in (19e–g). Interestingly, it is almost always the rhyme of the stressed syllable that deletes. There seems to be no overt exponence of the plural in these verbs other than the deletion process itself.

Finally, a very similar deletion process occurs in the formation of plural verbs in Alabama. This is shown in (20), taken from Hardy and Montler (1988a).

(20) Alabama—plural verb formation

	<i>Singular</i>	<i>Plural</i>	<i>Gloss</i>
a.	bala:-ka	bal-ka	'lie down'
b.	ibacasa:-li	ibacas-li	'join together'
c.	talbo:-li	talb-li [tal-li]	'make or build'
d.	batat-li	bat-li	'hit'
e.	kolof-fi	kol-fi [kol-li]	'cut'
f.	hałap-ka	hał-ka	'kick'
g.	cokkali-ka	cokka-ka	'go into'

The Alabama subtractive plural resembles that of Koasati in that the rhyme of the penultimate syllable deletes in (20a–f): a long vowel deletes in the first three words and the VC sequence of the rhyme deletes in (20d–f). However, in (20g) where the penultimate syllable is light (i.e. just a CV sequence), the entire syllable, rather than just the rhyme, deletes. The subtractive pattern for all the Alabama forms shown in (20) can be generalized in such a way that the last two positions of the penultimate syllable delete. The comparison of the subtraction patterns among these three languages highlights the particular challenges that piece-based theories would face in expressing morphological exponence when the process is neither additive nor feature changing. That is, there is no overt evidence for morphological exponence as “piece.”¹⁰ In process-based approaches, in contrast, the deletion itself would be the expression of the morphological exponence.

12.3.2 Augmentative Morphology

In augmentative morphology, exponence is expressed by the addition of a segment that is underspecified for phonemic content. It typically entails the addition of a moraic unit to the base that can be realized as vowel lengthening or consonant insertion among other possibilities. The specific phonological realization of the mora varies depending on the nature of the base. The process of adjectival emphasis in Shizuoka Japanese in (21) discussed by Davis and Ueda (2002, 2006) and by Trommer and Zimmermann (2010) serves as a simple example. The data are taken from Davis and Ueda (2002).

¹⁰ Two recent attempts to account for subtractive morphology in a “piece”-based approach to morphology are Trommer and Zimmermann (2010) and Bye and Svenonius (2012). Trommer and Zimmermann analyze subtraction as an affixal mora that is not realized phonologically resulting in deletion. Bye and Svenonius (2012) view subtraction as the affixation of an underspecified root node. Crucially, these accounts are both framed within Optimality Theory so that the resulting subtraction is a consequence of the specific ranking of the relevant constraints.

(21) Emphatic adjectives in Shizuoka Japanese

Subclass 1

	<i>Adjective</i>	<i>Emphatic form</i>	<i>Gloss</i>
a.	hade	hande	'showy'
b.	ozoi	onzoi	'terrible'
c.	yowai	yonwai	'weak'
d.	hayai	hanyai	'fast'
e.	karai	kanrai	'spicy'
f.	nagai	nan'ai	'long'
g.	kanafii	kannafii	'sad'
h.	amai	ammai	'sweet'

Subclass 2

	<i>Adjective</i>	<i>Emphatic form</i>	<i>Gloss</i>
a.	katai	kattai	'hard'
b.	osoi	ossoi	'slow'
c.	takai	takkai	'high'
d.	atsui	attsui	'hot'
e.	kitanai	kittanai	'dirty'
f.	kusai	kussai	'stinky'

Subclass 3

	<i>Adjective</i>	<i>Emphatic form</i>	<i>Gloss</i>
a.	zonzai	zoonzai	'impolite'
b.	kandarui	kaandarui	'languid'
c.	onzokutai	oonzokutai	'ugly'
d.	suppai	suuppai	'sour'
e.	okkanai	ookkanai	'scary'

A quick perusal of the data in (21) gives the impression that there is no uniform exponence between the three subclasses. In Subclass 1, the second syllable of the adjectival base begins with a voiced consonant, and a nasal consonant is inserted in the coda of the first syllable to form the emphatic. In Subclass 2, the second syllable of the adjectival base begins with a voiceless consonant, and the voiceless consonant is geminated. In Subclass 3, the first syllable of the adjectival base ends in a coda consonant, and the emphatic form of the adjective is formed by lengthening of the first vowel. Despite the superficial lack of uniformity, Davis and Ueda (2002) analyze the pattern as an instance of mora affixation, whereby the emphatic mora labeled μ_e is introduced as part of the affixation process. For instance, the underlying input for the emphatic adjectives for the first word in each of the three subclasses in (21) is given as in (22).

(22) Input to the emphatic adjective

- a. $/\mu_e + \text{hade}/$ b. $/\mu_e + \text{katai}/$ c. $/\mu_e + \text{zonzai}/$

As Davis and Ueda (2002) detail in their optimality-theoretic analysis, given the input forms in (22), the phonology alone determines the specific instantiation of the exponence as either nasal insertion (Subclass 1), gemination (Subclass 2), or vowel lengthening (Subclass 3). The generalization drawn from the analysis is that the exponence for the emphatic adjective is expressed as an additional mora in (22), and it is non-concatenative in the sense that it does not have consistent segmental content.

A similar type of mora augmentation process, although calculated from the right edge of the word, is found with the imperfective in Alabama (Hardy and Montler 1988b, Samek Lodovici 1992, Grimes 2002). The imperfective is illustrated in (23).

(23) Imperfective gemination in Alabama (. indicates syllable boundary)

	<i>Perfective</i>	<i>Imperfective</i>	<i>Gloss</i>
a.	ci.pii.la	cíp.pii.la	'small'
b.	ho.co.ba	hóc.co.ba	'big'
c.	mi.sii.li	mís.sii.li	'close eyes'
d.	a.taa.nap.li	a.tán.nap.li	'rancid'
e.	i.bak.pi.la	i.bak.píi.la	'turn upside down'
f.	i.si	íi.si	'catch'
g.	hof.na	hóof.na	'smell'
h.	is.ko	iis.ko	'drink'

The imperfective form is derived from the perfective by the addition of a moraic element. In (23a–d), the moraic augmentation is realized through gemination of the consonant at the beginning of the penultimate syllable, while in (23e–h) it is by lengthening the penultimate vowel. The difference between these two types of augmentation is phonologically determined: gemination occurs if the antepenultimate syllable is open as in (23a–d); otherwise, penultimate lengthening occurs. Shizuoka Japanese and Alabama, thus, share the same mechanism that exponence is expressed moraicly.

12.3.3 Autosegmental Affixation

Autosegmental affixation is instantiated by a wide variety of phenomena wherein distinctive features are utilized to express exponence as in consonant mutation or vowel change such as umlaut, or in tonal morphology where exponence is expressed by a certain tone or tone pattern. While typical cases of feature change include consonant mutation and umlaut in which only one element of the base is changed, others include morphologically-triggered harmony processes that cause a feature change on more than one element of the base. Discussion of autosegmental affixation phenomena is widely available in the literature, including Lieber (1987), Spencer (1998), Wolf (2007), Finley (2009), Akinlabi (2011), among others. It is often the case that specific featural and tonal alternations co-occur with particular affixes so that the featural or tonal change is not the sole exponence of the

morpheme. It will be shown below, however, that the featural change can be the only indication of exponence and as such, they are more strictly non-concatenative.¹¹

Initial consonant mutation is observed with the transitive–intransitive verb pairs in Nivkh (also known as Gilyak), a language isolate of Siberia. Data are from Spencer (1991: 19).

(24) Nivkh verbs (the final /d/ has a palatal pronunciation)

	<i>Transitive</i>	<i>Intransitive</i>	<i>Gloss</i>
a.	rʌŋzʌʌd	tʌŋzʌʌd	‘weigh’
b.	χavud	q ^h avud	‘warm up’
c.	yesqod	kesqod	‘burn something/oneself’
d.	vakzd	pakzd	‘lose/get lost’
e.	ɾad	t ^h ad	‘bake’

Intransitive verbs begin with a voiceless stop and transitive verbs with a continuant. The continuant is voiceless when the intransitive form begins with an aspirated stop (24b, e); otherwise, it is voiced. This suggests that the morphological exponence that marks transitivity is not a phoneme size unit but a subsegmental feature such as [+continuant] or [–continuant]. In an autosegmental analysis these features may be represented as floating in an underlying representation: [+continuant] for the transitive forms and [–continuant] for the intransitive forms, which are realized on the initial phoneme of the verb base. (See Lieber (1987), Wolf (2007), Finely (2009), and Bye and Svenonius (2012) for detailed analyses of other cases of initial mutation.)

The Ethiopian Semitic language Chaha presents an example of final consonant mutation. As discussed in McCarthy (1983, 1986), Rose (1994, 1997), and Banksira (2000), Ethiopian Semitic languages are characterized by the use of palatalization and labialization to mark various morphological categories. Consider the Chaha data in (25), taken from Kenstowicz (1994: 443–5), focusing on the exponence for the feminine imperative forms in the middle column.

(25) Chaha imperatives, second person singular (an apostrophe indicates an ejective)

	<i>Masculine</i>	<i>Feminine</i>	<i>Gloss</i>
a.	nəməd	nəmədʸ	‘love’
b.	nəqətʼ	nəqətʸ	‘kick’
c.	nəqəs	nəqəsʸ	‘bite’
d.	gəræz	gəræzʸ	‘be old’
e.	wətʼæq	wətʼæqʸ	‘fall’

¹¹ Consequently, we will not discuss some of the more well-known mutations systems such as that found in Fula or in some of the Celtic languages.

f.	nəqəb	nəqəb	'find'
g.	bəkər	bəkər	'lack'

The examples in (25a–e) show that the feminine form of the imperative is marked by the palatalization of the final consonant. This can be analyzed with the palatal feature [–back] as the exponence for the feminine imperative. Note that the forms in (25f, g) fail to undergo palatalization: (25f) ends in a labial consonant, and (25g), in a rhotic consonant. Chaha has phonemic palatalized consonants, while lacking palatalized labial and rhotic phonemes. Thus, the [–back] subsegmental feature that marks the feminine form of the imperative is constrained regarding the type of consonant with which it can associate. The language does not allow palatalized labials and rhotics, and consequently there is no distinct exponence for the feminine forms in (25f, g).

Contrastive with final consonant mutation, mobile mutation in Chaha, by way of morphological labialization, is extensively used in both the nominal and verbal morphology (see Banksira 2000). The data in (26), taken from Kenstowicz (1994: 443), demonstrate the perfective verb form with a third person masculine object. The forms in the left-hand column are unmarked for person.

(26) Chaha perfective (with third person masculine object)

	<i>he Verb-ed</i>	<i>he Verb-ed him</i>	<i>Gloss</i>
a.	nækæb	nækæb ^w	'find'
b.	dænæg	dænæg ^w	'hit'
c.	nædæf	nædæf ^w	'sting'
d.	nækæs	næk ^w æs	'bite'
e.	kæfæt	kæf ^w æt	'open'
f.	mæsær	m ^w æsær	'seem'
g.	qæt'ær	q ^w æt'ær	'kill'
h.	sædæd	sædæd	'chase'

The exponence for the third person masculine object is labialization, which can be represented by the feature [+labial]. Chaha has labialized consonant phonemes, but a coronal consonant cannot be labialized. The forms in (26a–c) show that the [+labial] feature, which marks the object, goes on the rightmost consonant. This is clearest in (26a): it is the final consonant that is labialized, not the preceding velar consonant. Examples (26d–h) are interesting because the final consonant is a coronal and cannot be labialized. These data illustrate that the [+labial] feature marking the object is mobile. The feature goes with the rightmost labializeable consonant, which is the second consonant in (26d–e) and the initial consonant in (26f, g). In (26h) all the consonants are coronal and cannot be labialized; there is no distinct exponence in this example. Thus, although

Chaha exhibits final consonant mutation as both palatalization in (25) and labialization in (26), the processes are contrastive in that labialization is mobile while palatalization is not.

Labialization and palatalization in Chaha not only occur independently of each other but can occur simultaneously to indicate a morphological class. For example, the impersonal form of the verb is marked by both labialization and palatalization subject to the constraints shown above in (25–6). Consider the impersonal forms in (27), taken from Kenstowicz (1994: 443) and Banksira (2000: 207).

(27) Chaha impersonal (forms have no overt person marking)¹²

	<i>Personal</i>	<i>Impersonal</i>	<i>Gloss</i>
a.	kæfæt	kæf ^w æt ^y	‘open’
b.	nækæs	næk ^w æs ^y	‘bite’
c.	bænær	b ^w ænær	‘demolish’
d.	nækæb	nækæb ^w	‘find’
e.	girəz	g ^w irəz ^y	‘Age!’
f.	t’as	t’as ^y	‘Infringe!’
g.	nitir	nitir	‘Separate!’ (from the teats)

In (27a, b, e), if the last consonant can be palatalized and one of the prior consonants can be labialized, then both labialization and palatalization occur. In (27c, d, g), the labial feature is mobile whereas the palatal feature is not; otherwise, (27g) would be realized as *[nit^yir]. Finally, (27f) shows that palatalization can still occur even if there is no eligible consonant for labialization. Thus, the Chaha impersonal in (27) offers an intricate case where a single morphological category is marked by two distinct subsegmental features.¹³

We now briefly discuss three cases of morphological exponence that is marked by a feature change on a vowel. Probably the most commonly cited example of this type is umlaut or fronting of a base vowel to mark plurals in German. Umlaut can be the only indicator of the plural in German although it is frequently accompanied by a suffix. Some examples are in (28), given in German orthography.

¹² The transcription of the palatalized consonants follows Kenstowicz (1994) where a superscript ^y indicates palatalization. For the final consonant of (27e) and (27f), Banksira (2000) would transcribe these as palatoalveolar fricatives.

¹³ While the discussion of the diachrony of the Chaha impersonal is beyond the scope of this chapter, if one compares the Chaha impersonal to the Arabic passive shown earlier in (1), the source for the labialization and palatalization that occur with the impersonal may be in the full vowels /u/ and /i/, assuming that the Chaha impersonal is cognate with the Arabic passive.

(28) German plurals marked by umlaut

	<i>Singular</i>	<i>Plural</i>	<i>Gloss</i>
a.	Garten	Gärten	'garden'
b.	Bruder	Brüder	'brother'
c.	Vogel	Vögel	'bird'
d.	Faden	Fäden	'thread'
e.	Tochter	Töchter	'daughter'
f.	Schnabel	Schnäbel	'beak'

The effect of umlaut is to change a base [+back] vowel to [–back]. In these examples, only the first of these two vowels change: the second vowel is already [–back]. The German umlaut process can be analyzed as invoking a floating [–back] feature or by a morphologically triggered vowel fronting rule.

A second example involving a change in a vowel feature is the Javanese elative, as discussed by Dudas (1975) and Wolf (2007). Relevant data are given in (29).

(29) Javanese elative

	<i>Plain</i>	<i>Elative</i>	<i>Gloss</i>
a.	alus	alus	'refined, smooth'
b.	aṅel	aṅil	'hard, difficult'

The elative form is distinct from its plain counterpart by the presence of tensing of the last vowel. This occurs even though it is not common for tense vowels to appear in closed syllables in Javanese.

While the subsegmental feature in the Javanese and German examples above affects one segment, a class of plurals in the Berber language Tamashek (also known as Tuareg, a Berber language spoken in Mali) demonstrates that more than one vowel of the base may be affected. The plural is marked by an ablaut pattern in which each of the vowels of the base changes in a different way, leaving unchanged the consonants and prosodic shape of the base. Consider the data in (30) from Heath (2005) as presented in Bye and Svenonius (2012).

(30) Tamashek plural class

	<i>Singular</i>	<i>Plural</i>	<i>Gloss</i>
a.	ǎ-dádis	i-dúdas	'small dune'
b.	ǎ-mágʷor	i-múgʷar	'large quadruped'
c.	e-féyer	i-fúyar	'bustard'
d.	t-ə-yúbbe	t-i-yubba	'gulp'
e.	ǎ-kárfu	i-kúrfa	'rope'
f.	ǎ-káfər	i-kúfar	'non-Muslim'

Focusing on the base vowels (not the prefixal ones) in (30), we notice that regardless of the quality of the vowels of the singular base, the first and second vowels of the base in the plural are always [u] and [a], respectively. It is the presence of these two vowels that marks the exponence of the plural. The Tamashek phenomena in (30) instantiate what has been termed “melodic overwriting” in the literature on non-concatenative morphology. (See Ussishkin (1999), Nevins (2005), Zimmermann and Trommer (2011) and Bye and Svenonius (2012) for details of issues concerning formal analyses.)

The Tamashek data in (30) is an example of where the vowels of the base change in different ways to mark a morphological class. More common are cases in which the vowels of the base all change in the same way to mark a morphological class: that is, the same subsegmental feature is realized on more than one vowel of the base. Such cases resemble harmony processes, but they are nevertheless morphological since the subsegmental feature expresses the exponence of the particular category. An example comes from the difference between completive and incompletive verbs in Kanembu (Nilo-Saharan) in (31), as is discussed by Akinlabi (1994) and Finley (2009).

- (31) Completeve—Incompletive alternations in Kanembu (tones are not indicated)

	<i>Completive</i>	<i>Incompletive</i>	<i>Gloss</i>
a.	gɔnəkɪ	gonɒki	‘I took / I am taking’
b.	dalləkɪ	dallɒki	‘I got up / I am getting up’
c.	barɛnəkɪ	barenɒki	‘I cultivated / I am cultivating’

Akinlabi (1994) demonstrates that the vowels of the completive and incompletive forms of the Kanembu verb reflect feature harmony. In the incompletive form in the middle column of (31), all the vowels are made with advanced tongue root: that is, they all bear the feature [+ATR]. The vowels of the completive in the first column of (31) are all made with a retracted tongue root: they share the feature [−ATR]. It is suggested that this example is different from German umlaut or the Javanese elative where the subsegmental feature marking the exponence shows up on only one vowel. It is also different from Tamashek, in which base vowels undergo different changes. In Kanembu, all vowels of the base form have the same subsegmental tongue root feature to mark the morphological class.

Additionally, it should be noted that morphological harmony processes may affect consonants and vowels together. A well-cited example of this is found in the first person marking in Terena, an Arawakan language of Brazil (Bendor-Samuel 1960, Akinlabi 1996, Finley 2009). Consider (32), cited from Akinlabi (2011). The base forms are in the first column and the first person forms are in the third column.

- (32) Terena—1st person forms

	<i>Base form</i>	<i>Gloss</i>	<i>First person</i>	<i>Gloss</i>
a.	ayo	‘(his) brother’	ãyō	‘my brother’

b.	arine	‘sickness’	ãrĩnẽ	‘my sickness’
c.	owoku	‘(his) house’	õwõ ^g gu	‘my house’
d.	nokone	‘need’	nõ ^g gone	‘I need’
e.	taki	‘arm’	ⁿ daki	‘my arm’
f.	piho	‘(he) went’	^m biho	‘I went’

In Terena, the first person forms are marked by nasalizing the phonemes of the base. The basic pattern displayed in (32) is that if the initial phoneme of the base is an obstruent, that consonant is prenasalized, as in (32e, f). If the first phoneme of the base is not an obstruent, it and all the subsequent sounds of the word become nasalized up to the first obstruent of the word. The consequence of the pattern is that the nasal exponence of the first person may be expressed on all phonemes of the base form, as in (32a, b), all phonemes up to the first obstruent, as in (32c, d), or just as prenasalization on the first consonant (32e, f). The Terena data in (32) have characteristics of nasal harmony systems (e.g. obstruents blocking the harmony, Walker 2011). However, the comparison of the Terena data in the left-hand column with those of the first person column indicates that the “nasal harmony” is a morphologically triggered marking for the first person, and therefore, is different from phonological nasal harmony in which the trigger is any nasal consonant.

Finally, we show that morphological exponence can be expressed solely by the use of tone. Given our definition of non-concatenative morphology, tonal morphemes fall under its purview since such morphemes are not associated with any consistent phonemic sequence. Most tonal morphemes that have been reported come from African languages. Consider, for example, the Benue-Congo language Tiv, discussed by Pulleyblank (1986) as well by Spencer (1991), where various verb tenses are indicated uniquely by specific tones. The examples in (33), taken from Spencer (1991: 163), illustrate the recent past forms. The first example is a verb lexically specified for an initial high tone; the other two possess an initial low.

- (33) Tiv—recent past (acute accent represents high tone; grave accent represents low tone)
- | | | |
|----|---------|------------|
| a. | yévèsè | ‘fled’ |
| b. | vèndé | ‘refused’ |
| c. | ngòhórò | ‘accepted’ |

According to Pulleyblank (1986), the high tone on the second syllable corresponds to the exponence of the recent past. This analysis is supported by the same verb forms in the general past tense, shown in (34). Data follow the presentation of Spencer (1991: 165), where “!” indicates a downstepped high tone, that is, a high tone that is somewhat lowered.

- (34) Tiv—general past
- | | | |
|----|---------|------------|
| a. | !yévèsè | ‘fled’ |
| b. | vèndè | ‘refused’ |
| c. | ngòhòrò | ‘accepted’ |

The general past form is marked by a low tone that is realized on the first syllable. If the first syllable has a high tone, as in (34a), it is phonetically realized as a downstepped high tone.

While the use of tonal morphemes is pervasive in African languages, it does also occur in tone languages of Asia, including Cantonese as discussed by Yu (2007c). In Cantonese, verbs that have underlying level tone can be nominalized with the use of rising tone, although the productivity of the process is not fully investigated. Examples in (35) are taken from Yu (2007c: 191). (The numbers next to the items indicate tone levels, e.g. “1 1” indicates a low level tone, “3 3” a mid level tone, and “3 5” a mid rising tone.)

(35) Cantonese nominalization

	<i>Verb form</i>	<i>Nominalization</i>
a.	sou 3 3 ‘to sweep’	sou 3 5 ‘a broom’
b.	jɛu 1 1 ‘to grease’	jɛu 3 5 ‘oil’
c.	wɔ 2 2 ‘to listen’	wɔ 3 5 ‘an utterance’
d.	liu 1 1 ‘to provoke’	liu 3 5 ‘a stir’
e.	tan 2 2 ‘to pluck’	tan 3 5 ‘a missile’

The nominalizing morpheme in these examples is clearly not segmental but a mid rising tone pattern. Yu (2007c) mentions other possible morphological uses of the mid rising tone in Cantonese, and discusses previous autosegmental analyses where a mid tone and a high tone occur on the tonal tier as the representation of the morphological exponence of the nominalization pattern.

Although less common, we also find languages in which stress or pitch-accent is used to mark morphological exponence. English verb-noun pairs like *contrást–cóntrast* and *impórt–ímpor*t are sometimes mentioned as stress being used as a derivational device (e.g. Spencer 1991: 16), but the status of these English pairs is highly controversial (see Trommer (2012) and Bermúdez-Otero (2012) for recent discussion). Hidasta (Siouan), a pitch-accent language, invokes the morphological use of pitch for the derivation of its vocative form, as recently described by Park (2012). All of the vocative forms in this language have falling pitch on the last syllable; no other changes are required except for the lengthening of the last vowel, if it is short, so as to be able to carry the falling pitch. Some examples from Park (2012: 356–7) are given in (36).

(36) Hidasta vocative (acute accent represents high tone; ‘`’ represents a falling tone)

	<i>Base form</i>	<i>Vocative</i>	<i>Gloss</i>
a.	marisá	marisáà	‘my son’
b.	masáàwi	masaawíi	‘my aunt (father’s sister)’
c.	magúù	magúù	‘my grandmother’
d.	masígisa	masigisáà	‘my brother-in-law (women’s brother-in-law)’

Regardless of the location of high pitch in the base form, the vocative is indicated in a uniform way, with a falling pitch on the final vowel. The example in (36c) shows that if the base form already has falling pitch on the final vowel, the vocative form will take the identical pitch. The pattern illustrated by the vocative is unusual even in Hidasta: while it is not uncommon for the pitch pattern of the base form to change under affixation, the vocative is the only example where pitch is the lone exponence of a morphological class.

12.4 SUMMARY AND CONCLUSION

In this chapter we have given a typological characterization of non-concatenative (non-reduplicative) morphological phenomena. We framed our discussion with a specific focus on the expression of exponence, and categorized non-concatenative morphology into two distinct types: templatic and a-templatic. Templatic morphology involves morphological restrictions on the shape of words. We further divided templatic morphology into two types based on exponence. In particular, Semitic languages demonstrate that the template itself can be the unique exponence of a category. A second more common type involves a concatenative affix that imposes a templatic subcategorization requirement on the base to which it attaches. With respect to a-templatic non-concatenative morphology, a wide range of phenomena was surveyed: they included subtractive morphology, moraic augmentation, and autosegmental affixation.

While our primary goal has been to give a descriptive perusal and offer a typological characterization of a variety of non-concatenative phenomena, by no means does the descriptive nature of our overview undermine the significance of specific theoretical issues and controversies pertinent to non-concatenative morphology found in the literature. One such issue is how to handle non-concatenative phenomena in a “morphology-as-pieces” approach. Particularly problematic would be cases of morphological subtraction where the process of deletion itself seems to mark the category. A morphology-as-pieces approach would view non-concatenative morphology as epiphenomenal: that is, it results from the nature of representation that can include underspecified “pieces” as well as incorporating Optimality Theory where non-concatenative effects can arise as a consequence of the specific constraint ranking. Bye and Svenonius (2012) arguably give the most detailed account of this approach to non-concatenative morphology. Others, such as Downing (2006), have independently argued against prosodically defined templates in morphology. However, the existence of non-concatenative morphology as a distinct phenomenon is relatively unproblematic for a construction-based view of morphology along the lines of Booij (2010). Constructional schema involve form–meaning pairings, and a template or a subsegmental feature would just be part of the form in the form–meaning pairing. Tsujimura and Davis (2011a, 2011b) give an explicit illustration of how aspects of non-concatenative morphology can be analyzed in a construction grammar approach.

Another issues related to non-concatenative morphology that we have not discussed in our overview is whether phenomena like word shortening (e.g. *fridge* from *refrigerator*) and word blends (e.g. *brunch* from *breakfast* and *lunch*) should fall under the purview of non-concatenative morphology. On the one hand, the output of these word formation processes can often be defined by prosodic templates. On the other hand, a number of researchers have argued that the arbitrariness of the phenomena is not something that a grammar should account for. Alber and Arndt-Lappe (2012) provide an overview discussion on this issue.

A final matter concerns how the often murky distinction between inflectional and derivational morphology relates to non-concatenative morphology. Non-concatenative processes are often found with types of morphology that are normally considered inflectional. Nonetheless, non-concatenative phenomena frequently exhibit characteristics of derivational morphology even when expressing inflectional type categories. As we have contrasted in (17) and (18) above, for instance, the distinct exponence of the subtractive morphology of the Tohono O'Odham perfective verb is only found with those verbs ending in consonants and not in vowels. While tense/aspect marking is typically considered inflectional, the type of phonological restriction found in the Tohono O'Odham perfective is often more characteristic of a derivational process. Consequently, the critical criteria for distinguishing derivation from inflection will include the question of their applicability to non-concatenative morphology and whether such morphology is invariably derivational or can be both derivational and inflectional. This further touches on the issue of whether there is a difference between templatic and a-templatic non-concatenative morphology. We leave these challenging questions for future research.

ACKNOWLEDGMENT

We would like to thank Sabine Arndt-Lappe, Laura Downing, Andrew Koontz-Garboden, Tracy Alan Hall, Natalie Operstein, Jochen Trommer, Adam Ussishkin, Rachel Walker, and the two editors of the volume for valuable input to this chapter.

CHAPTER 13

ALLOMORPHY

MARY PASTER

THIS chapter deals with allomorphy, which we will define as a situation in which a single lexical item, meaning, function, or morphosyntactic category has two or more different phonological realizations depending on context. As we will see, the contexts that condition allomorphy may involve phonological, morphosyntactic, and/or lexical factors. Of particular interest is the fact that there is evidence for two different types of phonologically conditioned allomorphy; this poses a significant analytical challenge. Allomorphy is observed in both inflectional and derivational affixes, as well as in roots. Given the theme of the volume, this chapter will focus on examples of allomorphy in derivational affixes.

There is not complete agreement among researchers regarding what constitutes allomorphy or how to differentiate the types of allomorphy. One major point of contention concerns a possible distinction between “suppletive allomorphy” and “rival affixes,” as will be discussed. In this chapter I will take the position that there is no such distinction, but it should be understood that there exist other views on this issue.

13.1 TYPES OF ALLOMORPHY

Allomorphy can be divided into two main types. The first is predictable phonological (non-suppletive) allomorphy. This describes a common situation in which a productive phonological rule applies to an affix in some contexts but not others, yielding multiple different surface forms of the affix. A simple example is found in Luganda, where (as in many Bantu languages) there is vowel harmony resulting in alternations in some derivational suffixes. The applicative suffix, for example, has the surface forms [ir] and [er], as shown in (1).

(1)	oku-gul-ir-a	‘to buy for’	oku-som-er-a	‘to read for’
	oku-zin-ir-a	‘to dance for’	oku-kol-er-a	‘to make for’
	o-ku-sal-ir-a	‘to cut for’		

The generalization here is that the suffix surfaces as [er] if the preceding vowel is mid; otherwise [ir]. It makes sense to assume there is a single underlying form of the affix, namely /-ir/, and that there is a rule changing the high suffix vowel to mid when preceded by a mid vowel, because we find that this is a general process in the language. We observe the same pattern, for example, in the causative suffix, as shown in (2).

- (2) o-ku-yimb-is-a 'to make sing' o-ku-som-es-a 'to teach (make read)'
 o-ku-kub-is-a 'to make beat' o-ku-koz-es-a 'to make make'

A traditional analysis of this situation, in a model of grammar where morphology feeds phonology, would be that a single underlying form of the affix attaches to every stem. After the affix is attached to a particular stem, the stem + affix combination is passed to the phonology. The phonology scans the stem + affix unit to determine whether the context is present for the application of the relevant phonological rule (in this case, one changing a high vowel to mid when preceded by a mid vowel). In some cases the context is present, so the rule applies; in other cases the context is not present, so the rule does not apply. Since the segment that changes belongs to the affix, the result is that the affix has two different surface forms depending on a phonological property of the stem to which it attaches.

This type of allomorphy is relatively straightforward, though as we will discuss later, there are some complications (primarily involving distinguishing this type of allomorphy from the other type of phonologically conditioned allomorphy to be discussed shortly). It should also be noted that some researchers do not use the term "allomorphy" for the situation I have just described; they might just call this "morphophonology" and use the term "allomorphy" only to refer to suppletive allomorphy—to which we now turn.

Suppletive allomorphy¹ (sometimes just called "allomorphy") is the second type of allomorphy, and the one we will focus on most closely in the remainder of the chapter. The fundamental property of suppletive allomorphy is that it always involves two or more different underlying forms of the relevant affix or root. This can be contrasted with the regular phonological allomorphy discussed above, where a single underlying form has multiple surface realizations (though, as we will see, it is possible and even common for both types of allomorphy to occur in the same affix).

I define suppletive allomorphy as any situation where the same set of morphosyntactic/semantic features is expressed by two or more surface forms in complementary distribution that have different underlying forms. However, other researchers have a more

¹ The term "suppletive," pronounced [sə'plɪrɪv], refers to the fact that the two (or more) different forms "supplement" each other as in geometry, where supplementary angles add up to 180 degrees. A better term might have been "complementary allomorphy" since we have the term "complementary distribution" to describe the relationship between allophones of a phoneme—but "suppletive allomorphy" (or simply "suppletion") is widely used.

restrictive definition of allomorphy that in various ways requires the underlying forms to be phonetically similar. For example, Stockwell and Minkova's (2001: 73) definition of "allomorphy" requires the allomorphs to have a "historically valid" relationship—that is, to be etymologically related. In this type of framework, cases where the underlying forms are not phonetically similar would constitute "suppletion" (but not "suppletive allomorphy") in the case of roots, or "rival affixes" in the case of affixes. Under the view that I will advance in this chapter, there is no difference between "rival affixes" vs. "suppletive allomorphy." Since suppletive allomorphs have separate underlying forms and are selected within the morphology (rather than the phonology), the degree of phonetic similarity or difference is not referred to by the grammar. The allomorphs may be very phonetically similar, reflecting a shared etymology (but for whatever reason, not being relatable to a single underlying form) or they may be extremely different, suggesting that they were historically completely distinct items that were collapsed into a suppletive relationship over time, as with *go* ~ *went* in English. From the perspective of phonology, once it is determined that the allomorphy is suppletive rather than regular (phonological), phonetic similarity between the allomorphs is irrelevant. Adopting this position, henceforth in this chapter I will include in the discussion of "suppletive allomorphy" some examples that might be called "rival affixes" by other researchers.

Suppletive allomorphy may be conditioned by at least three different factors (sometimes a combination of multiple factors): phonological, morphosyntactic, and lexical. Examples of each of these will be given below. We begin with the phonological type.

Phonologically conditioned suppletive allomorphy (or PCSA) is a phenomenon where multiple surface forms are conditioned by a phonological factor, but (crucially) not via the operation of a phonological rule on a single underlying form. Rather, it is the distribution of two or more separate underlying forms that is phonologically determined—often in a way that does not relate at all to the phonological shape of the affix or root in question. An example is found in Dutch, where two suffixes that are commonly used to derive adjectives from nouns are *-isch* /is/ and *-ief* /iv/ (Booij and Lieber 1993). For nouns ending in *ie* /i/, the distribution of the adjectival suffix allomorphs is determined by the stress pattern of the stem, as follows: *-isch* is used if the stem has final stress, while *-ief* is used if the stem-final syllable is unstressed. Examples are shown below (Booij and Lieber 1993: 25).²

(3) a.	sociologie	'sociology'	sociolog-isch	'sociological'
	blasfemie	'blasphemy'	blasfem-isch	'blasphemous'
	allergie	'allergy'	allerg-isch	'allergic'
b.	preventie	'prevention'	prevent-ief	'preventive'
	constrúctie	'construction'	construct-ief	'constructive'
	integrátie	'integration'	integrat-ief	'integrating'

² Booij and Lieber (1993) do not mark stress on the derived forms. Based on Booij (2002: 114), it may be inferred that forms with the *-isch* suffix assign stress to the last stressable syllable before the suffix, but this is a generalization referring to the "native" suffix *-isch*, while the *-isch* that attaches to non-native

Notice that there is no plausible phonological rule that will convert *-isch* to *-ief* (or vice versa) in the relevant context; the rule would have to change the word-final fricative based on the input stress pattern of the stem and it would have to be specific to this particular suffix. Therefore this is best analyzed as a case involving two separate underlying forms for the suffix (i.e. suppletive allomorphy).

In many cases it is a bit more difficult to determine whether a particular case involves phonological (predictable) allomorphy vs. PCSA. Kiparsky (1996: 17) gives a set of criteria distinguishing between the two:

(4)	<i>Allomorphy</i>	<i>Morphophonology</i>
a.	item-specific	general (not item-specific)
b.	may involve more than one segment	involve a single segment
c.	obey morphological locality conditions	observe phonological locality conditions
d.	ordered prior to all morphophonemic rules	follow all morpholexical processes

By Kiparsky's own admission, the convergence of his criteria is only "fairly consistent," and these criteria "cannot claim to provide an automatic resolution of every problematic borderline case" (1996: 16). Criterion (4a), involving the generality of the pattern, is perhaps one of the more useful criteria. If a phonological rule/constraint proposed to account for a pattern of allomorphy in a particular morpheme can also account for one or more other patterns of allomorphy in the same language, this suggests that the allomorphy is best analyzed as resulting from the application of phonological rules/constraints to a single underlying form (i.e. that the allomorphy is not suppletive). If, on the other hand, the rule/constraint that would need to be posited to account for a pattern of allomorphy would only be manifested in that particular morpheme, then the allomorphy is more likely to be suppletive. However, in some cases it can be argued that a particular affix or group of affixes is associated with a "co-phonology" (Inkelas et al. 1997, Inkelas 1998) that might result in particular rules or constraint rankings that do not apply throughout the entire language. This is especially useful if a construction that includes a particular affix (or one of a group of affixes) seems always to have a particular phonological property (as with, for example, stress-shifting affixes).

A related factor (hinted at in the discussion of the Dutch example) that bears on whether to analyze a particular pattern as suppletion or item-specific rule application is the plausibility of the proposed rule. Suppose that a particular morpheme presents

stems is listed separately (Booij 2002: 76) as a non-native suffix. Therefore, Booij may be assuming there are two *-isch* suffixes (though it is not clear how one would distinguish them). Stress is marked on one form with *-ief*: *cònservatief* 'conservative' (Booij 2002: 106), suggesting that *-ief* gets the main stress in words where it occurs, but there are not enough examples to demonstrate that this generalization holds for all forms with *-ief*.

the only instance in the language of the phonological configuration that would trigger the rule. In this case, the rule would be both “general” (in the sense that it applies everywhere in which the phonological environment for its application is met) and “item-specific” (since it would only apply to one morpheme). Therefore, in this hypothetical situation, criterion (4a) is of no help in determining whether this is a case of rule-driven allomorphy or suppletion. A secondary consideration in such a situation is plausibility. If the rule is item-specific but is also formally simple, then this is an argument in favor of rule-derived rather than suppletive allomorphy. On the other hand, if the proposed rule would be formally complex (perhaps involving multiple operations or affecting multiple segments simultaneously—thus relating to Kiparsky’s criterion (4b)), the pattern should be analyzed as suppletive. Applying this criterion does, of course, require the researcher to commit to some formal model for which it is clear what constitutes an allowable operation, trigger, target, etc., so that the plausibility of a rule can be assessed.

In summary, though some criteria can be established for identifying suppletive vs. non-suppletive allomorphy, there still exists a substantial gray area. This is not surprising, since many examples of suppletive allomorphy probably result from historical processes of the restriction of productive phonological rules to particular morphological contexts. Thus, they may have many of the properties of regular phonological processes but also lack some of those properties.

Having considered phonological conditions on suppletive allomorphy, we move now to a second possible type of conditioning factor in suppletion, namely morphosyntactic context. A particular allomorph of an affix may occur only in the presence (or the absence) of another affix. For example, McPherson and Paster (2009) argued that in Luganda, the causative suffix has an allomorph /-iz/ that occurs when preceded by the applicative suffix /-ir/. Recall from (2) (repeated in (5)) that the usual form of the causative suffix is /-is/ (alternating between surface variants [-is] and [-es] based on vowel harmony, as discussed earlier).

- (5) o-ku-yimb-is-a ‘to make sing’ o-ku-som-es-a ‘to teach (make read)’
 o-ku-kub-is-a ‘to make beat’ o-ku-koz-es-a ‘to make make’

When it co-occurs with the applicative /-ir/, however, the causative surfaces as [-iz], as shown in (6). There is nothing about the phonological environment created by the addition of /-ir/ that would trigger a change from /s/ to [z] in the causative suffix, so this should be treated as suppletive allomorphy.

- (6) ba-ji-tu-mu-fumb-ir-iz-a
 3SG.SUBJ-9.OBJ-1PL.OBJ-3PL.OBJ-COOK-APPL-CAUS-FV
 ‘they make us cook it for her’

An interesting question that arises here is whether it is the actual affix that conditions the allomorphy, or the morphosyntactic features associated with the affix. In this particular example, is it the feature [+applicative] that conditions the use of the /-iz/ form of

the causative, or is the allomorphy triggered by the /-ir/ suffix that expresses the applicative? In this case it is not possible to distinguish empirically between the two options since there is only one form of the applicative suffix. However, in principle, if there were multiple suppletive allomorphs of the applicative, we could test whether all of the allomorphs patterned together in conditioning the use of /-iz/ for the causative. We will discuss this issue further in Section 13.4 in connection with the question of whether allomorphy sheds light on the (non-)existence of the “morpheme” as a meaningful unit.

A third and final type of condition on suppletive allomorph distribution is the lexical type. Certain stems condition the use of specific affix allomorphs on a completely arbitrary basis that must be lexically specified. For example, in Polish, the passive participle of verbs is formed differently depending on the lexical class of the verb. Some examples are given in (7) (Swan 2002: 303–4; the passive forms shown here are masculine singular forms).

(7) a.	psuć	‘spoil’	psuty	‘spoiled’
	dąć	‘puff’	dęty	‘puffed’
	osiągnąć	‘attain’	osiągnięty	‘attained’
b.	nieść	‘carry’	niesiony	‘carried’
	wieźć	‘transport’	wieziony	‘transported’
	piec	‘bake’	pieczony	‘baked’
c.	kupić	‘buy’	kupiony	‘bought’
	tworzyć	‘create’	tworzony	‘created’
d.	писаć	‘write’	pisany	‘written’
	widzieć	‘see’	widziany	‘seen’

According to Swan (2002: 302–4), passive participles are formed as follows. “First conjugation” verbs (examples in (7a)) take the passive marker *-t-* (with verbs ending in *-qć* having *q* changed to *ę* and verbs ending in *-nqć* having their *n* palatalized). “First conjugation obstruent consonant-stems” (those whose infinitives end in *-ść*, *-źć*, or *-c*) form their passives with *-en-* (which changes to *-on-* except when followed by the masculine plural ending *-i*), as shown in the examples in (7b). Second conjugation verbs (examples given in (7c), with infinitives ending in *-ić* or *-yć*) also take *-en-/on-*, but this is added to a different stem form from that of the examples in (7b). Finally, verbs with infinitives ending in *-ać* or *-eć* form the passive with *-n-* (and the *e* of infinitives in *-eć* changes to *a*), as shown in (7d). Notice that some of the generalizations here are stated in terms of the phonological form of the stem, but these do not form natural classes, and there is no way to derive the passive suffix forms based on the phonological shape of the stem. Therefore, the passive endings must be assigned based on the (arbitrary) lexical class of the verb.

A question raised by examples of this type is whether they are best analyzed as being productively formed, as opposed to being lexically stored (as a root + affix unit). If the lexical entry for the bare root already specifies which allomorph must be used for an affix that may attach to it, then perhaps the root + affix combination has a separate lexical representation of its own. To some extent this is an empirical issue to be decided on a case-by-case basis. Some considerations when determining whether or not the root +

affix combination is listed in the lexicon for a particular case would include (1) whether the affix in question is used productively elsewhere in the language (if not, as with English *oxen*, then the combination is more likely lexically listed), and (2) whether other affixes may intervene between the root and the relevant affix (if so, then the combination should probably not be lexically listed).

Processing studies involving the speed of retrieval may also shed light on this question. Some studies have defended a view where regular forms are formed productively while irregulars are stored lexically (e.g. the dual route model of Pinker 1991, 1997). Others have argued that even irregulars are productively formed in the grammar (e.g. Levelt et al. 1999). However, some have proposed based on frequency effects that both regular and irregular forms may be stored (e.g. Baayen et al. 1997a, 2002). Discussions of predictions for regular vs. irregular forms often assume that there is only one regular pattern; cases of suppletion where more than one allomorph is fully productive (*modulo* the contextual restrictions giving rise to the allomorphy) and therefore arguably regular present interesting variations on the notion of regularity.

Summarizing the examples of allomorphy that we have seen in this section, we have seen examples of what I have described as the two main types of allomorphy—predictable phonological allomorphy and suppletive allomorphy—and within what I have called suppletive allomorphy, we have seen examples of three subtypes based on the type of conditioning involved. We can schematize this conception of allomorphy as in Figure 13.1.

One caveat here is that, as has already been hinted at above, the different types and subtypes of allomorphy are not mutually exclusive; a given example may include multiple types of allomorphy affecting a single affix.

An example showing both phonological (predictable) and suppletive allomorphy is found in Turkish (Lewis 1967). In Turkish, the causative is marked by /-t/ with polysyllabic stems ending in vowels, /r/, or /l/; the suffix /-DIr/ (where *D* is a coronal stop and *I* is a high vowel) is used with all other stems (except for some specific monosyllabic stems that take a different, lexically determined allomorph). Some examples are shown below (Haig 2004).

(8)	bekle-t-	‘wait-CAUS’	öl-dür-	‘die-CAUS’
	bayıl-t-	‘faint-CAUS’	ye-dir-	‘eat-CAUS’
	getir-t-	‘bring-CAUS’	çalış-tır-	‘work-CAUS’

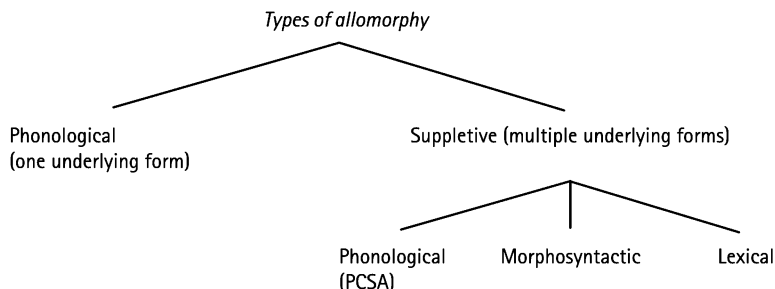


FIGURE 13.1 Types of allomorphy

Notice here that in addition to the pattern of PCSA that determines which form of the suffix will be selected, we also observe predictable phonological allomorphy in the “elsewhere” form of the suffix. The initial consonant of the suffix varies between [t] and [d] due to assimilation to the voicing of the preceding segment, and the vowel of the suffix varies between [i] and [ü] due to harmony with the stem vowels. In a linear derivational model of grammar, the analysis would be that the morphology first selects the appropriate underlying form of the suffix (in this case based on phonological factors), and then the regular phonology applies after the suffix is attached.

We also find cases where two or more different types of conditions are referenced in a single instance of suppletive allomorphy. For example, in Russian (Timberlake 2004), the reflexive marker exhibits both morphologically and phonologically conditioned suppletive allomorphy. The reflexive suffix has two allomorphs, [s'a] and [s']. The [s'a] variant is always used in active participle forms. In other forms, the [s'a] variant occurs after consonants, while the [s'] variant occurs after vowels (Timberlake 2004: 345). The examples below are from Wade (2002: 137) (transliterations are mine).

(9)	akupaju-s'	'I bathe myself'	onkupajet-s'a	'he bathes himself'
	kupajt'e-s'	'bathe yourselves!'	kupaj-s'a	'bathe yourself!'
	onakupala-s'	'she bathed herself'	onkupal-s'a	'he bathed himself'

Though the allomorphs are phonetically similar to each other, I treat the phonological conditioning as PCSA since the pattern does not result from the application of any general rule of the language. Thus, the distribution of the /s'/ form is restricted both phonologically and morphosyntactically: it occurs only after a vowel-final verb that is not in an active participle form. It may initially seem surprising that the distribution of a single allomorph may be restricted by a combination of phonological and morphosyntactic conditions, but really this is no different from the unremarkable fact that any given affix (including those exhibiting PCSA) will generally also attach only to stems of a particular syntactic category. For example, in the Turkish case discussed earlier, in addition to exhibiting PCSA, the causative suffix is limited to attaching to verbs. We tend to ignore these types of restrictions since they are more mundane or trivial than what we would call morphologically conditioned suppletive allomorphy. In fact, however, they can be explained by the same mechanism—namely, affixes selecting (“subcategorizing”) for particular morphosyntactic features. Whether or not a particular case of selection will result in suppletive allomorphy depends only on whether another less restricted form of the affix exists to fill in the gap left by the selectional requirements of the restricted affix. We will discuss the relationship between blocking/gaps and suppletive allomorphy further in Section 13.3.

13.2 ANALYZING SUPPLETIVE ALLOMORPHY

There are a number of different ways of analyzing suppletive allomorphy; some alternatives will be presented in Section 13.5 in connection with outstanding theoretical issues in

the analysis of suppletion. The approach I will assume here is a subcategorization-based model (see also Lieber 1980, Kiparsky 1982b, Selkirk 1982, Yu 2003, 2007a, Paster 2006, 2009). In this approach, suppletive allomorphy results when two or more different affixes with the same meaning have different subcategorizational requirements, which are selectional requirements imposed by affixes on stems. Affixation satisfies missing elements that are required as specified in the lexical entry for each morpheme. The grammar will attempt to use the most specific affix available to express a given set of morphosyntactic/semantic features, so the most restricted allomorph of a morpheme will be tried first. If its subcategorizational requirements are not met, a less restricted allomorph (typically the “elsewhere” form) of the morpheme is used. Thus, the distribution of affixes in this approach is determined by the subcategorizational requirements. These are properties of each affix, and they determine the types of stems to which each affix will be allowed to attach.

An important feature of this approach is that subcategorizational requirements may include morphosyntactic, semantic, and/or phonological properties of the stem. These are all assumed to be part of the makeup of the underlying form of the morph, so that satisfaction of all three types of requirements is done as part of the process of affixation—that is, within the morphological component of the grammar. Note that this feature of the model is what theoretically differentiates PCSA from predictable phonological allomorphy, since the former is handled within the morphology (along with all other types of suppletive allomorphy) while the latter is done in the phonology.

By now the relation between suppletive allomorphy and morphological gaps in this approach may be apparent. In the following section we discuss gaps in more detail.

13.3 ALLOMORPHY AND THE INFLECTION VS. DERIVATION DIVIDE

Cases of gaps in derivational morphology are well known; one such example comes from English:

(10)	brighten	*dullen
	darken	*dimmen
	blacken	*brownen
	whiten	*greyen
	redde	*bluen
	thicken	*thinne
	shorten	*tallen

The generalization observed here is that the *-en* suffix will only attach to an adjective if it ends in an obstruent (there are further restrictions as well). Notice that when the suffix

cannot attach to a particular stem, the result is a gap. There simply does not exist a morphological way to derive the relevant form of “blue” in English; this meaning can only be conveyed periphrastically, for example as “make blue”.³

In a morphological subcategorization-based approach to allomorphy, this situation is treated just like a case of PCSA: the *-en* affix subcategorizes for an adjective ending in an obstruent, so it will not attach to any stem that does not meet this requirement. The difference between this and PCSA is simply that in PCSA, there would be a second, “elsewhere” allomorph that could convey the relevant meaning when the stem cannot take *-en*. Because there is no such allomorph in English, the result is a gap, so *-en* alternates not with another affix but with the syntactic construction “make X”.

It was once thought (e.g. by Aronoff 1976, cited in Carstairs 1988) that derivational morphology never exhibits suppletive allomorphy because the concept of suppletion relies on the existence of a paradigm, and derivational morphology does not involve paradigms. Carstairs (1988: 74) argued against this idea based on the simple fact that there are attested examples of suppletive allomorphy in derivation (as we have seen above). Further, it has been argued that derivational morphology is, in fact, paradigmatic (see, e.g., van Marle 1985, 1986).

Carstairs makes the generalization (1988: 75) that “most” of the examples in his sample of fifteen cases (from ten different languages) involved inflection rather than derivation. Though this may have been a premature generalization given the small number of languages considered in that study, it does appear to be upheld in Paster’s (2006) survey of 137 examples of PCSA (from sixty-seven languages), though neither survey claims to be balanced or representative. Carstairs (1987) proposes an explanation for his generalization based on a principle of inflectional parsimony, which states that for any combination of morphosyntactic features that can mark members of a particular word class, each word will have exactly one inflectional realization. Essentially, this is “blocking”—the existence of one surface realization of a set of features precludes the formation of all other realizations. According to Carstairs (1988: 76), there are three ways in which a language can resolve a potential violation of parsimony (from a diachronic point of view). First, all but one of the possible realizations could drop out of use. Second, all of the realizations could be distributed arbitrarily into conjugations or declensions. And third, all of the realizations could be distributed according to some independent principle (whether semantic, syntactic, morphological, or phonological). Thus, the development of PCSA is just one of many ways in which a language can adhere to the parsimony principle.

Carstairs goes on to argue that

a principle of parsimonious coverage does appear to exercise an influence over not only inflectional morphology and syntactic structure but also certain areas of lexical

³ An affix could be used to express the meaning “make blue,” but only in what I would deem a “creative” use of the relevant affix. For example, one might say *blue-ize* or *blue-ify* or even *em-blue-en*, but most native speakers would probably judge these words not to be grammatical, strictly speaking; they sound (at least to this native speaker) like intentionally humorous coinages rather than fully acceptable native words.

organisation involving even monomorphemic items. If this is so, it would be surprising if the principle could not also affect morphologically complex lexical items, including derived words. (Carstairs 1988: 79)

An example of the former type of effect in English is pointed out, where animal species have one item in each of the categories adult male, adult female, and young. According to Carstairs, it is this type of lexical organization or derivation, which he describes as “meaning-driven,” in which parsimonious coverage is apparent. Another example is male vs. female titles of ranks of the British peerage (*duke* and *duchess*, etc.). An example of English derivation that is *not* meaning-driven would be deverbalization using *-ion*, *-al*, *-ment*, *-ance*, or stress shift. In this type of derivation, the exact meaning of the noun derived from the verb is not predictable (e.g. *remit* vs. *remission*), and we do not find parsimonious coverage: some verbs have multiple possible nominal forms (e.g. *commis-sion*, *commital*, *commitment*). In Carstairs’ view, the fact that only meaning-driven derivation obeys the principle of parsimonious coverage accounts for why cases of PCSA are more often inflectional than derivational. Carstairs ultimately suggests that a distinction between “meaning-driven” and “expression-driven” morphology may turn out to be more useful than inflection vs. derivation.

13.4 MORPHEMES AND MORPHOMES

Kiparsky (1996) contrasts two competing approaches to suppletive allomorphy: selection vs. replacement. Under a selectional approach (of which the subcategorization-based approach outlined in Section 13.2 is one example) a morpheme is a set of alternants, some of which occur only in certain contexts, and one of which is considered the “elsewhere” morph. Kiparsky points out that the concept of the “morpheme” is unnecessary in this approach, though it can be defined as “a set of morphs in a blocking relationship” (1996: 18). To this definition one might add that the morphs must have identical semantic features, since there are cases of blocking relations among sets of morphs that one would probably not wish to call a morpheme.⁴

In the replacement approach, on the other hand, the concept of the morpheme is essential. Each morpheme has a single underlying phonological form, and allomorphy rules replace this form with other allomorphs in certain contexts. Kiparsky notes that the selectional approach, but not the replacement approach, predicts that morphological conditioning can be triggered by specific morphs and by morphological categories, but not by “morphemes.” Kiparsky argues that the empirical facts support the selection approach in a number of ways. In many cases we find morphological gaps, which should

⁴ For example, in Nimboran (Inkelas 1993, based on Anceaux 1965), the plural object marker blocks the dual subject marker.

not exist under the replacement approach;⁵ in some cases we find optionality, where two morphs overlap in their distribution, which again is not predicted by the replacement approach; allomorphs are not “outwardly sensitive”; and finally we do not find allomorph selection based on derived phonological properties,⁶ nor do we find phonological processes triggered by a “basic” allomorph which is later replaced via an allomorphy rule.

Thus, following Kiparsky’s arguments, we may conclude that what we know about allomorphy does not require or support the concept of the “morpheme” as anything more than a descriptive device. It is useful for analysts to be able to talk about relations among different affixes that indicate the same meaning or morphosyntactic category in different contexts, but the grammar does not necessarily need to treat those affixes as belonging to an abstract morpheme that represents the meaning or category. A type of evidence from allomorphy that could support the concept of the morpheme would be a case of morphosyntactically conditioned suppletive allomorphy where the conditioning affix itself has multiple allomorphs. If all of the allomorphs of the conditioning morpheme (morpheme A) patterned together in conditioning allomorphy in the other affix (belonging to morpheme B), this would suggest that the morphosyntactic/semantic feature bundle of morpheme A (i.e. the morpheme itself) is responsible for conditioning the allomorphy and the grammar would have to be able to refer to the morpheme. If, on the other hand, only one allomorph of morpheme A ever triggers a pattern of allomorphy in morpheme B, this would suggest that the grammar does not need to be able to refer to the abstract morpheme.

A related question is whether we need to be able to reference the “morpheme” in describing allomorphy. The morphome, proposed by Aronoff (1994), is purely functional and even a step more abstract than the morpheme, since it does not require semantic coherence. A test case for the necessity of the morphome to analyzing allomorphy might be one similar to the hypothetical situation discussed just above, where multiple allomorphs of morpheme A trigger allomorphy in morpheme B. In order to make the case for a morphomic level of representation, we would need to be able to show that the set of triggers is a natural class morphologically—but not semantically or phonologically. At present I am not aware of any such cases; this remains an open empirical question.

⁵ This argument rests on the assumption that there is a difference between a productive affix that leaves a morphological gap vs. one that is simply not fully productive. The assumption seems to be valid, since there are attested cases of gaps where the construction in question is otherwise highly productive (e.g. the past participle form of the verb “stride” in English).

⁶ It is not exactly true that allomorphy is never conditioned by derived phonological properties. However, I know of no example that cannot be accounted for in a cyclic model; that is, allomorphy can be sensitive to a phonological property that was derived on an earlier cycle. A potential example of sensitivity to a derived phonological property is the Dutch case described earlier involving PCSA conditioned by stress, since regular stress is not typically assumed to be present in underlying representations (and is therefore taken to be derived).

13.5 OTHER THEORETICAL ISSUES

There remain a number of other theoretical issues involving suppletive allomorphy that are still under debate. One of the most interesting issues involves the directionality of conditioning. In the subcategorization-based approach, word-building is assumed to proceed from the inside out—starting with a root and adding successive layers of affixation (the natural way of thinking about this is derivationally, but the notion of successive layers does not necessarily preclude a parallel model). An inside-out approach to word-building entails that allomorphy may be triggered only by an “inner” morph (the root or an affix closer to the root); allomorphy may not be triggered by a peripheral element (i.e. an affix that is farther away from the root).

The question of directionality of conditioning in suppletive allomorphy has not to my knowledge been systematically studied cross-linguistically on a large scale. However, in Paster’s (2006) cross-linguistic study of PCSA, no absolutely convincing cases of “outside-in” conditioning were found; some marginal examples were discussed and shown to be reanalyzable. The lack of examples of this type was used as an argument for the subcategorization-based approach and against the then-standard approach to PCSA in Optimality Theory (OT), in which PCSA and predictable phonological allomorphy are handled within the same component of the grammar (McCarthy and Prince 1993a, b). In this approach to PCSA, all possible allomorphs are listed in the input, and inter-ranked phonological and morphological constraints select the optimal surface allomorph. As discussed by Paster (2006, 2009), this approach predicts rampant outside-in conditioning of suppletive allomorphy because all parts of a word are present in the input and assembled in parallel. Thus, the OT approach seems to overgenerate. Later research, however, has uncovered further possible examples of outside-in conditioning. To the extent that these examples hold up and can be shown to require an analysis in terms of outside-in conditioning, these examples support alternative approaches to allomorphy—or perhaps modifications to the present version of the subcategorization-based model.

Wolf (to appear) discusses two examples of apparent outside-in conditioning in PCSA involving inflectional affixes in Armenian and Kayardild. In the Armenian case, it is claimed that a plural suffix is added to some non-plural forms in order to create a minimal disyllabic stem needed for the attachment of the plural possessive suffix. Some examples are given in (11) (Wolf to appear: 8–9, citing data from Vaux 2003).⁷

⁷ I have attempted to use examples of derivational rather than inflectional morphology in this chapter given the theme of the volume; however, in this case as well as the case from Embick (2010) to be discussed shortly, it is necessary to consider some examples from inflection. To the extent that these examples hold up, there is no particular reason to believe that derivational morphology could not behave the same way.

(11)		‘cow’	‘cows’	‘cat’	‘cats’
	‘X’	gov	gov-ər	gadu	gadu-nər
	‘my X’	gov-əs	gov-ər-əs	gadu-s	gadu-nər-əs
	‘your (sg.) X’	gov-ət ^h	gov-ər-ət ^h	gadu-t ^h	gadu-nər-ət ^h
	‘his/her/its X’	gov-ə	gov-ər-ə	gadu-n	gadu-nər-ə
	‘our X’	gov-ər-ni-s	gov-ər-ni-s	gadu-ni-s	gadu-nər-ni-s
	‘your (pl.) X’	gov-ər-ni-t ^h	gov-ər-ni-t ^h	gadu-ni-t ^h	gadu-nər-ni-t ^h
	‘their X’	gov-ər-ni-n	gov-ər-ni-n	gadu-ni-n	gadu-nər-ni-n

The issue here is that the forms in the bottom left-hand corner (‘our cow,’ ‘your (pl.) cow,’ and ‘their cow’) appear to have a plural suffix (-ər) on the noun, whose presence is conditioned by the outer suffix -ni. Thus, on the surface, this looks like outside-in conditioning.

Additionally, Wolf discusses a number of cases of apparent PCSA in stems, conditioned by affixes. These cases provide arguments against the subcategorization-based approach and in favor of Optimal Interleaving (Wolf 2008), a model based on OT but incorporating serialism in a way that, Wolf argues, eliminates the overgeneration problem encountered by the traditional OT approach to PCSA.

Embick (2010) proposes a model that also allows for some outside-in conditioning, but in more limited instances. Embick’s model, termed C₁-LIN (‘cyclicity-linearity’) is based on a version of Distributed Morphology (DM; Halle and Marantz 1993; Embick and Marantz 2008). In Embick’s approach, as in DM generally, the phonological content of an affix only becomes visible at Vocabulary Insertion, which applies during spell-out. Spell-out is done cyclically in phases. Therefore, allomorph selection is sensitive to phonological properties of items that are being spelled out on the same cycle. This means that some outside-in conditioning of suppletive allomorphy is predicted since multiple adjacent items may be spelled out in a cycle, but otherwise no outside-in effects are predicted. Embick (2010: 61) presents a possible case of outside-in conditioning from Hupa (Golla 1970), reproduced in (12).⁸

- (12) a. no:xoWtiW
 no- x^wi- W- t- tiW
 ADV OBJ 1SG TRANS put
 ‘I put him down.’
- b. na:se:ya?
 na- si- e- ya?
 ADV PERF 1SG go
 ‘I have gone about.’

⁸ Golla (1970: 31) identifies /W/ as a voiceless rounded glottal fricative.

- c. siWda
 si- W- da
 PERF 1SG sit
 'I am sitting.'

The generalization is that 1sg subject agreement prefix *e-* is used if it is preceded by perfective prefix and verb is non-stative; the prefix *W-* occurs elsewhere. If this example holds up, it constitutes outside-in conditioning since the triggering morpheme (the perfective prefix) is a peripheral affix, occurring farther from the root than the prefix that exhibits the allomorphy.

Related to directionality is the question of whether a property of one word can trigger allomorphy in another word. The short answer may seem to be yes, given that some examples are discussed in the literature. However, given that the theories discussed here do not deal especially well with such cases, and given the apparent reanalyzability of the known cases, it is possible that the longer answer may turn out to be no.

One obvious and well known example of word-external conditioning is the *a/an* alternation in English. However, given the status of *a/an* as a clitic, which we take to mean that it is part of the same phonological word as the morpheme that triggers the allomorphy, this is not a very good example of word-external conditioning.

Another possible case is found in Mafa (Chadic, Cameroon). Le Bleis and Barreteau (1987) report that the verbal suffix indicating "le directionnel de rapprochement" occurs as *-ká* when preceding a word that begins with a consonant, and as *-káďá* elsewhere (Le Bleis and Barreteau 1987: 108–9; English translations mine).

- (13) á mbálə-ká kəda
 il-INACC chasser-RAPPR chien
 'il court après le chien vers nous'
 ['he runs after the dog towards us']
- m bálə-ká yim (áduwzlak)
 (no interlinear glosses provided)
 'íl a puisé de l'eau (l'a versée dans la jarre) et l'a rapportée'
 ['he drew some water (poured it in the jar) and brought it back']
- n táv-káďa aa gírzhe
 il-ACC monter-RAPPR sur rocher
 'íl est monté sur le rocher (qui se trouve entre l'endroit d'où il vient et celui où se trouve le locateur)'
 ['he climbed onto the rock (which is located between the location from which he came and the location of the speaker)']
- kalədə-káďá
 tomber.CAUS-RAPPR
 'jette-le vers moi!'
 ['throw it towards me!']

We do not have evidence for the directional morpheme in Mafa being a clitic (which might have allowed us to say that the following word is its host within a single phonological word, as in the English *a/an* example). The directional morpheme is not described as having any other properties that we could use to argue that it is a clitic rather than an affix; for example, it seems always to occur immediately after the verb stem rather than allowing other words to intervene. It also does not seem plausible to reanalyze this as predictable phonological allomorphy, since the two allomorphs differ in two segments. We can tentatively conclude that the Mafa case is a legitimate example of word-external conditioning. Future research may reveal more cases in other languages that would bolster our confidence in the Mafa example and prompt some revisions to current theories that would allow us to model the phenomenon more straightforwardly.

CHAPTER 14

NOMINAL DERIVATION

ARTEMIS ALEXIADOU

14.1 INTRODUCTION

NOMINAL derivation (henceforth nominalization) is a process that derives a noun from another word category, normally a verb or an adjective. Thus it is a category changing operation which can take place with or without inducing a change on the form of the source element. Across languages, both morphological types of nominalizations are possible. For instance, in English we have, on the one hand, nominals derived from, for example, verbs via the addition of a derivational affix as well as so-called zero derived nominals that lack any overt morphological change. Characteristic examples are given in (1):

(1)	<i>Verb</i>	<i>Noun</i>
a.	to find	the finder, the finding
b.	to jump	the jump

The main concern of this chapter will be cases of nominalization that involve category changing morphology.

The literature on nominalizations has focused on various aspects that make this process so interesting for linguists from very different perspectives. Clearly, the obvious difference between the nouns and the verbs in (1a) has to do with the fact that the nominalization externally behaves as a noun, as it can occupy an argument position in its own right. Consider the set of examples in (2), from Baker and Vinokurova (2009). *Finding* in (2b) and *finder* in (2c) both contain a verbal root and a nominalization affix, both occupy NP positions (the subject of the clause), both are introduced by the determiner *the* and contain an object introduced by *of*. In contrast, *finding* in (2d) appears more verbal. There is no determiner present, and the object of

the nominal bears accusative Case, although the nominal is morphologically identical to that in (2b):

- (2) a. Chris found my wallet in the stairwell.
- b. The finding of the wallet took all afternoon.
- c. The finder of the wallet returned it to the front desk.
- d. Finding my wallet so quickly was a big relief.

The nominals in (2) thus show a mixed-categorical behavior (nominal and verbal) to a varying degree. Several authors have tried to come up with explanations to account for, on the one hand, the semantic similarity and morphological relationship between the verb and the nominals that can be derived from it and, on the other hand, the fact that simply the noun is not quite like the verb in a number of ways, and that there is a gradience when it comes to verbal as opposed to nominal behavior (Ross 1972). A recent overview of this discussion is offered in Alexiadou (2010a, b) and I will not go into that here.

In this chapter, I will look in some detail at the syntax and semantics of mainly deverbal nominalizations from the perspective of polysemy (though at the end of this chapter I will briefly look at de-adjectival nominalizations). While inheritance of the argument structure (AS) of their related verb clearly disambiguates between, for example, event and other readings, as well as between state and quality readings of de-adjectival nominalizations (Grimshaw 1990, AS being only possible with the former, see Sections 14.3, 14.6, and 14.7), still the question arises why do nominalizations surface with the same form, although they differ in meaning (see Beard 1990)? Can we establish any generalizations as to which meanings will cluster together? Does identity in form imply identity in meaning (iconicity principle)? In an attempt to answer these questions, I will examine participant nominalizations, deverbal nominalizations that are ambiguous between event and result/object readings, and de-adjectival nominalizations that are ambiguous between a state and a quality reading.

The chapter is organized as follows: I will first discuss so-called participants nominalizations, focusing on *-er* and *-ee* nominals in English. I will then turn to deverbal nominals which are ambiguous between event and object/result readings, for example English *-ing* and *-ation* nominals. In both domains, I will have nothing to say about AS inheritance (see the aforementioned sources). In Section 14.7, I will briefly look at de-adjectival nominalizations, as these have been argued to parallel the behavior of deverbal nominalizations in the sense that they also split into two classes, and AS inheritances is again the factor that disambiguates between the two. In my discussion, I will leave out nominal derivations from other nouns, for example English *childhood*, *lordship*, etc., as these do not offer such a parallel. The reader is referred to Lieber (2004) and Trips (2009) for detailed discussion and references.¹ In what follows, while English will be my point of departure, I will offer cross-linguistic remarks.

¹ The analysis offered in these works is in terms of lexical semantics. It is not clear to me how syntactic approaches to nominalization would treat such examples; perhaps they could be analyzed as cases of root compounding.

14.2 TYPES OF NOMINALIZATIONS

Comrie and Thompson (2007) identify two types of lexical nominalization, involving either actions/states or participants. In participant nominalization, the noun formed relates to a semantic role of the nominalized verb (agent, patient, instrument). In action/state nominalization, the noun formed refers to an action or a state. Both are classified as lexical nominalizations in opposition to syntactic nominalizations that involve relative clauses, but see Ntelitheos (2012) for arguments that participant nominalizations involve relative clause formation as well.

In a language like English, various affixes are used to form participant nominalizations, as shown in (2), for example *-er*. Affixes such as *-ing*, and *-ation*, are predominantly used to form action/state nominalizations. I will discuss these in detail here. I begin my discussion with participant nominalizations, concentrating on the behavior of *-er* and *-ee* nominals in English. These affixes are quite intriguing for the following reasons: first of all, in English, at first sight they seem to have a complementary distribution: *-er* is predominantly involved in the formation of nominals relating to the external argument of the nominalized verb, while *-ee* is predominantly involved in the formation of nominals relating to the complement of the nominalized verb. Both statements will be refined in the next subsections. This refinement will bring us to the second interesting property these affixes share: they are highly polysemous, that is, they can form both subject-oriented as well as object-oriented nouns, raising the obvious question how should morphological theory capture their distribution. A similar concern is raised for *-ing* and *-ation* nominals. Third, while across languages we find the counterpart of *-er* nominals, this is not the case for *-ee* nominals, a point already raised in Booij and Lieber (2004), and more recently in Štekauer et al. (2012). Finally, and most importantly, these as well as *-ing* and *-ation* nominals and de-adjectival nominalizations have been analyzed from a variety of perspectives (syntax, AS, lexical semantics, and cognitive linguistics), providing us with a fertile ground on which to compare approaches.

14.3 ER NOMINALIZATIONS

14.3.1 Different Types of *-er* Nominalizations

As stated in Booij and Lieber (2004), the morphological literature on *-er* nominals has established that these show a wide variety of meanings, see the examples in (3) and (4), from Booij and Lieber (2004):

(3) a. subject-oriented *-er*

<i>base</i>	<i>theta-role of subject</i>	<i>derived noun</i>
write	agent	writer
drive		driver
open	instrument	opener
print		printer
hear	experiencer	hearer
please	stimulus	pleaser

b. object-oriented *-er*

<i>base verb</i>	<i>thematic role</i>	<i>derived noun</i>
fry	patient/ theme	fryer
keep		keeper
dine	location	diner 'place where one dines'
sleep		sleeper 'train in which one sleeps'

As is well known, there is a large number of forms derived with *-er* that have non-verbal bases, illustrated in (4), from Booij and Lieber (2004):

(4) non-deverbal *-er*

English		
<i>base</i>	<i>base category</i>	<i>derived noun</i>
London	noun	Londoner
Village		villager
Five	measure	fiver

Schäfer (2011) identifies a further type of *-er* nominalization, which is, however, restricted to German, so called event *-er* nominalizations, see (5) from Schäfer (2011).

(5) ein Piepser

- a. a beeper (an agent who beeps)
- b. a beep (a/one beeping event)

Schäfer notes that the formation of event denoting *-er* nominals is not an idiosyncratic phenomenon restricted to a small number of verbs. Rather it is very productive within the class of verbs that can be classified as semelfactives (e.g. *cough*, *beep*, *knock*, etc.). Ryder (1999) observes that event *-er* nominals are also possible in English, albeit not subject to the same restrictions identified for German by Schäfer, for example *breather*, *no-brainer*.

It is often stated that there is a difference in productivity between subject denoting *-er* nominalizations, on the one hand, and object-denoting *-er* nominals, on the other. For instance, Schäfer (2011), and references therein, notes: "while virtually every verb

projecting an external argument allows a *-er* nominal denoting the external argument, only a small subset of verbs allows *-er* nominals to denote the internal argument (object).” It has also been observed that non-verbal *-er* derivations are not fully productive in English. This means that we cannot use any adjective, preposition, or noun to form a corresponding *-er* nominal. However, this should not be taken to mean that there are no interesting generalizations to be made about what kind of non-verb derived *-er* nominals there are. On the contrary, as Schäfer (2011) signals, noun derived *-er* nominals are clearly restricted: there are noun classes that allow *-er* formation relatively productively, while others that do not allow it so productively (e.g. animals: ?*dogger*). The *-er* nominals from nouns denoting places (cities, villages, countries, and so on) denote people who live at this place and can be formed relatively productively. This suggests that formation of these nouns is productive within a particular domain. Thus, any morphological analysis should be able to offer an account of the patterns found and also explain the difference in productivity between subject-oriented and non-subject-oriented *-er* nominals. As will be discussed in Section 14.3.3, there is also a cross-linguistic difference when it comes to non-subject-oriented *-er* nominals. For instance, such forms are absent from German, but are present in English and Dutch. Our analysis should be able to capture this as well.

Let me now examine the two groups, subject *-er* and non-subject *-er* nominals, in turn.

14.3.2 Subject *-er* nominals

There is some consensus in the literature that English subject *-er* nominals can be divided into two major sub-classes (see Rappaport Hovav and Levin 1992, Fabb 1984, Keyser and Roeper 1984, van Hout and Roeper 1998, to mention a few), the relevant semantic property being whether they refer to an actual event or not. That is, *-er* nominals vary with respect to the [$\pm$ event] specification. It has also been pointed out that [+event] *-er* nominals correspond to the external argument of the base verb irrespective of the thematic role that this verb assigns to its external argument (agent, causer, holder, experiencer, instrument). I call this the “external argument generalization.”² Thus they are not necessarily agentive, see also Booij and Lieber (2004), see (3a).

[−event] *-er* nominals also fall into two thematic groups. In the first group, we find [+agentive] nouns, as in (6), in the second group, we find [+instrumental] *-er* nominals, as in (7). Both classes have in common that they denote entities which are *designated* for some specific job or function but which do not have to be actually involved in such a job or function (the [−event] property).

² A remark of clarification is in order here. The morphological literature classifies these forms in terms of subject- vs. object-oriented. The literature on argument structure uses the term external argument, which corresponds to a subset of the forms that can function as grammatical subjects, i.e. those that are in a sense deep subjects.

- (6) lifesaver, fire-fighter, teacher → a person educated for a specific job
 (7) a grinder → machine intended for grinding things

Rappaport Hovav and Levin (1992) establish the following correlations:

- (8) An instrumental reading is possible only for the nominals derived from verbs for which the expression of an instrumental performing a “subject” role is available.
 (9) AS is inherited by event *-er* nominals only.

Concentrating on (8), the external argument generalization is independent of AS inheritance. Compare the instrument in (10) with the instrument in (11). They differ in that the instrument in (10a) can occur as the subject of a corresponding sentence (11b) while this is not possible for the instrument in (11a) (see 11b).

- (10) a. Mary opened the can with the new gadget
 b. The new gadget opened the can
 (11) a. Bill ate the food with a fork
 b. *The fork ate the meat

Instruments of the former type are called *intermediary instrument*, instruments of the latter type are called *facilitating* or *enabling instruments*. They note that these two types of instruments differ in that only the former can be understood to perform the action expressed by the verb (to some extent) independently, a property that qualifies them as subjects of these verbs (see also Kamp and Rossdeutscher 1994, Alexiadou and Schäfer 2006 and references therein). Crucially, corresponding instrumental *-er* nominals are only possible for verbs that combine with intermediary instruments.

- (12) a. opener (agent or instrument)
 b. eater (agent but **not instrument**)³

In English, Spanish, German, and Dutch the same affix is used to form nouns denoting both an instrument and an agent. Other languages, however, seem to use distinct affixes. For instance, in Greek the affix that derives the agentive nominal is different from the affix that derives the instrument nominal. Agentive nouns are built on the basis of the affix *-tis*, while instrumental nouns are built on the basis of the affix *-tiras*, *-tiri*, *-tirio*, *-tra*. Note that, as Dressler (1986) observes, the instrumental (and locative, see below) affixes seem more complex than the agentive ones in the sense that they contain

³ Rochelle Lieber remarks that synthetic compounds of the type *odor-eater* are possible, meaning a kind of insole for shoes. Crucially, however, the instrumental interpretation is only available in the presence of the first member of the compound, which corresponds to the object of the verb *eat* as in *This insole eats odor*. In addition, the NP *the insole* corresponds to a primary instrument in the aforementioned example.

a further consonant, namely *-r-* (although in Classical Greek the affix used was more similar to that of instrumental nouns, for example *-tor*, the final *-r* disappeared from Modern Greek nominal declension):

- (13) a. *Base verb* *Noun (agentive)*
 pezo peh-TIS
 play-1SG player-MASC.
 litrono litro-TIS-MASC
 save-1SG saver
 b. *Base verb* *Noun (instrumental)*
 anigo anih-tir-i
 open-1SG opener-NEUT.

In Romance, for example French, two different morphemes are used for the formation of *-er* nominals, “*-eur*” and “*-oir(e)*” which are, however, etymologically derived from the same Latin root “*-or*,” see also Rainer (2005b) and references therein. Interestingly, the difference between the two is that “*-eur*” tends to specialize for external argument denoting nouns while “*-oir(e)*” forms nouns denoting locations and instruments (see Alexiadou and Schäfer 2010 for discussion and references).

The Germanic pattern is also found in languages outside of Indo-European, for instance in the Austronesian language family. Consider Table 14.1 illustrating Saisiyat nominalizations, taken from Yeh (2011), see also Ntelitheos (2007) on Malagasy. The formation of all the nouns shown in Table 14.1 involves the nominalizer *ka-* that combines with different Voice markers, a typical characteristic of this language family.

Table 14.1 Formation of argument nouns in Saisiyat

Type	Form	Example
Agentive	ka-ma-V	ka-ma-'omalop 'hunter'
		ka-ma-ka:at 'writer'
Patient	ka-V-en	ka-'alop-en 'game'
		ka-ka:at-en 'the thing to be written, homework'
	V -in-	'-in-alop 'game been hunted'
Locative	k-in-a:at	'what is written, book, letter, word'
	ka-V-an	ka-'alop-an 'hunting area'
	ka-ka:at-an	'place for writing, desk'
Instrument	ka-V	ka-ka:at 'pen'
	Ca~V	'a-'alop 'hunting instrument'

A similar situation is also observed in Hausa, where the affix *ma-* derives agentive, instrumental, and locative nouns (data from Štekauer et al. 2012: 171):

- (14) a. *maà-ikàc-i*
 AG-work-M.SG
 worker
 b. *ma-girbi*
 harvesting tool
 c. *majema*
 tannery

According to Štekauer et al. (2012), this type of polysemy is wide-spread. Other languages where such a polysemy is found include Hungarian, Finnish, Hindi, Indonesian, Spanish, Slovak, and Swedish.

14.3.3 Object *-er* Nominals

Two remarks are in order here. First, Booij and Lieber note that non-subject *-er* nominals almost always denote things rather than people. In addition, they denote the affected object not the effected object, see (15).

- (15) a. *baker* (a baked potato)
 b. *broiler* (a broiled chicken)

Second, while examples as in (15) can be found in both English and Dutch, their distribution is cross-linguistically restricted. For example, they occur in Hungarian, Slovak, Serbo-Croatian, and Chinese (see Štekauer et al. 2012), but they do not occur in German, the Romance languages, and Greek. As Štekauer et al. (2012: 176) state, “there are no instances in our sample of patient/instrument polysemy/homonymy with no agent in this relation. On the other hand, the patient’s absence in the one-to-many relation between agent/instrument is quite common”. Moreover, German lacks locative *-er* nominals such as *sleeper*. In Greek, and Romance, we do find locative nominals formed productively from verbs, but these are formed on the basis of an affix rather similar to the one used for instrumentals. This is illustrated with Greek examples in (16):

(16)	<i>Verb</i>	<i>Noun</i>	
	<i>dikazo</i>	<i>dikas-tirio</i>	cf. <i>dikas-TIS</i>
	<i>judge-1SG</i>	<i>court room-NEUT</i>	<i>judge-MASC.</i>
	<i>shediazō</i>	<i>shedias-tirio</i>	<i>shedias-TIS</i>
	<i>design-1SG</i>	<i>designing room-NEUT</i>	<i>designer-MASC.</i>

As mentioned, object *-er* nominals are not fully productive, thus it has been suggested that they are (in fact need to be) lexicalized. While it is indeed possible to find *-er* neologisms,⁴ I will assume here that most object *-er* nominals are lexicalized, it is then a different but important question why English and Dutch have more of these *-er* nominals than, for example, German or the Romance languages or Greek (see also the discussion in Booij and Lieber 2004).

14.4 TOWARDS AN ACCOUNT OF POLYSEMY

As is clear from Section 14.3, the type of affixal overlap under discussion is not restricted geographically or genetically. To the extent that *-er* and its counterpart in other languages create different semantic types of deverbal nouns, in principle an account in terms of polysemy seems plausible, and several such approaches have been proposed. A summary of these is found in the following sources. Booij (1986) identifies three different types of explanation for the polysemy of derived words, which are not in principle mutually exclusive. Rainer (2005b) also offers an overview of the approaches to this problem, including a diachronic–synchronic discussion of agentive and instrumental nouns.

According to Booij (1986), one option is to associate one and a very general meaning with the word-formation process. A second option is to assume one core or prototypical meaning and derive the other meanings by extension rules. A third option is to assume that polysemy reflects differences in the thematic grids of the verbal bases. The thematic grid of a verb provides information concerning the thematic roles associated with its arguments, that is, internal arguments/complements and subjects. Let me now offer a summary of the analyses various scholars pursued in the past.

Booij (1986) argues that *-er* binds the external argument of the verb. Since the external argument can bear a variety of thematic roles, the interpretation of the derived nominal will vary accordingly. Booij, however, argues that instrumental *-er* nouns do not have the same status as the other *-er* nominals. Therefore, he suggests deriving the instrumental interpretation by means of a conceptual extension schema that allows a shift from personal agent, through impersonal agent to instrument. Booij considers the instrument noun to be different because he observed that there are cases where an instrument *-er* nominal can be formed although the corresponding verb does not readily tolerate an instrument as its subject. For example, in Dutch the deverbal noun *smelter* ‘melter’ may be interpreted as an instrument, while the sentence *De warmte smelt het ijs* ‘The heat melts the ice’ is odd. Note, however, that Booij admits that this is ungrammatical in the absolute sense. Other authors have pointed out that Dutch differs from English when it

⁴ Consider, for instance, the following example provided by Rochelle Lieber from *Outdoor Life* 2005: “I had taken bears before and had been hunting for several years for a truly outstanding bear, and here one was standing broadside at 20 yards. I didn’t have to think twice about this bear. It was a **shooter**.”

comes to instruments as subjects. For instance, Guilfoyle (2000), following van Voorst (1996), judges (17a, b) ungrammatical:

- (17) a. *De sleutel opende de deur
 'The key opened the door'
 b. *De steen brak het raam
 'The stone broke the window'

Guilfoyle argues that a parameter exists that distinguishes between two types of languages: in languages of type A the external argument position is closely associated with the initiator of the event (Dutch), hence judgments are degraded; in languages of type B the external argument is associated with a participant in the event, and does not necessarily need to be an initiator (English). This would then predict that instrumental *-er* nominals should be less productive in Dutch. However, as Alexiadou and Schäfer (2006) argued, instrument subjects behave alike in all languages and are acceptable only under two conditions which force a Causer or an Agent interpretation of the instrument respectively. Under this view, it is expected that in both Dutch and English instrument *-er* nominals should be formed, and Booij's objection that led to the use extension schemata can be dealt with. In both languages, such formations obey the external argument generalization.

Rappaport Hovav and Levin (1992) also analyze *-er* at the level of AS: *-er* binds the external argument of the verb to which the affix attaches, and it can bear any of the roles that the external argument of the verb can have, for example agent or instrument. Their account also considers object denoting *-er* nominals. These authors note that nominals such as in *fryer* or *looker* have an interpretation that is close to the interpretation that the base verb receives in the middle construction. Thus, it was proposed that these nominals are in fact derived from the middle version of underlying verbs where the theme (the argument denoted by the *-er* nominals is the (allegedly base generated) external argument of the verb.

This analysis enables us to understand why such object denoting nominals are impossible in German, Romance, and Greek. Specifically, Alexiadou and Schäfer (2010) argue that a reason for this difference could be that English and Dutch form morphologically unmarked middles while German and Romance mark their middles with the reflexive pronoun "*sich/se*" (cf. Schäfer (2008) for a proposal which correlates this difference in morphological marking with a difference concerning the syntactic position of the theme in middles; in Dutch and English middles, the theme is a derived external argument, while in German/Romance middles, it remains in its VP-internal base position; Greek uses non-active morphology on its middles and thus is amenable to an analysis similar to German/Romance).

However, as Booij and Lieber (2004) point out, there are several cases that cannot be captured by this analysis. For example, the verb *keep* does not have a middle form, still, *keeper* is a patient nominal. Nor can locational nouns such *diner* be explained in this way. Also problematic are the non-verbal derived nominals such as *Londoner*. How could such forms be derived, if there is no corresponding verbal base?

Heyvaert (2010) attempts to keep the intuition expressed in Rappaport Hovav and Levin in order to account for all types of *-er* nominals, including locative ones, but not really non-deverbal ones. First, she points out that non-agentive *-er* nominalizations have much in common with the middle construction in that “they also designate entities of which the properties are conducive to a specific process and can also be analyzed as expressing a modal relationship between a process and an entity: a *cooker* is more than an *apple that cooks* and a *broiler* is more than a *chicken that broils*: they *cook and broil WELL due to their properties*”. Importantly, Heyvaert notes, their semantics can be further differentiated according to the categories that were also distinguished for middles. Consider (18), taken from Heyvaert (2010):

- (18) *bestseller*: facility/quality-oriented
 frontloader: destiny-oriented
 sleeper: feasibility and facility-oriented
 kneeler: destiny-oriented
 cooker: destiny- and result-oriented

A further point made by Heyvaert is that agentive lexicalized *-er* nominalizations also typically imply a dynamic type of modality: most prototypical agentives imply the dynamic modality of *ability* (*can*) (e.g. *teacher*). Some also imply regularity or persistent habit (*will*) (e.g. *drinker*). Instrumental *-er* nominals embody one of the fundamental choices that is offered by *-er* suffixation, that is, that between an agentive and a non-agentive one.

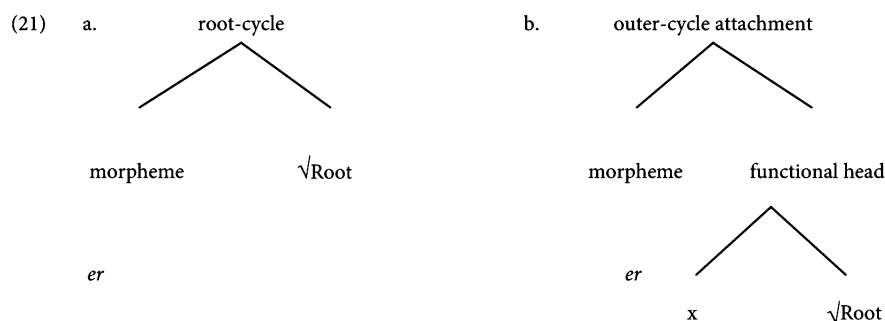
Instruments by definition hover between being able to carry out a process themselves (as agent-like participants) and letting others carry it out (as non-agentive entities). Those instrumental *-er* nominalizations that are non-agentive resemble middle constructions in that they profile an entity that has properties that let an implied agent perform a particular action (e.g. *stroller*). Agent-like instrumentals, on the other hand, foreground the agent-like ability of the tool which they refer to (e.g. *transmitter*, *toaster*). A large group of instrumental *-er* nouns lies in between the agentive and non-agentive type: depending on which perspective is chosen, they can be interpreted as either agentive or non-agentive. (Heyvaert 2010: 66)

An account in terms of lexical semantics is given in Booij and Lieber (2004), who propose that the affix *-er* forms a concrete noun, and the skeletal contribution of this affix will be nothing more than the features [+material, dynamic ([]), <base>]. Their analysis is cast within the model of lexical semantics proposed by Lieber (2004), according to which the lexical semantic representation of lexemes (and of affixes, at least to a certain extent) is composed of two parts: a semantic/grammatical *skeleton* and a semantic/pragmatic *body*. The distinction *skeleton-body* is roughly reminiscent of the distinction proposed in Rappaport Hovav and Levin (2010) between event structure template and root, and is therefore an optimal instrument for representing the lexical decomposition of verb meaning in a way that is comparable to syntactic approaches. Lieber’s model

- (19) In a configuration in which semantic skeletons are composed, co-index the highest non-head argument with the highest head argument. Coindexation must be consistent with semantic conditions on the head argument, if any.
- (20) writer [+material, dynamic ([i], [+dynamic ([i], []))]
 -er write

Turning now to syntactic approaches, Alexiadou and Schäfer (2010) offer a syntactic analysis developed within the distributed morphology (DM) framework, cf. van Hout and Roeper (1998), Baker and Vinokurova (2009), and Ntelitheos (2012) among others. The basic ingredients of this framework can be stated as follows (see Marantz 1997, 2001, Arad 2003): Language has atomic, non-decomposable and category-neutral elements, which we refer to as roots. Roots combine with features, the functional vocabulary, and

build larger elements. On this view, words are not primitives. The primitives of word formation are the roots and the functional vocabulary they combine with. Word categories are determined by category defining functional heads. Derivational endings are part of this functional vocabulary. Some words are built out of roots. Some others are built out of other words. This means that there are two cycles for word-formation (Marantz 2001), and distinct properties are associated with each one of them. From this perspective, affixes are underspecified as to their locus of insertion, that is, they can appear in structures that have distinct meaning:



Merger with root implies:

1. negotiated (apparently idiosyncratic) meaning of root in context of morpheme;
2. apparent semi-productivity (better with some roots than others);
3. meaning of construction cannot be an operation on “argument structure” but must depend on root semantics independent of argument structure (see Barker 1998);
4. corollary of the above: cannot involve the “external argument” of the verb.

Merger above a category-determining morpheme implies:

1. compositional meaning predicted from meaning of stem;
2. apparent complete productivity;
3. meaning of structure can involve apparent operation on argument-structure;
4. can involve the external argument of a verb.

As already mentioned, *-er* nominalizations differ in productivity and whether or not they can involve the external argument of the verb. Alexiadou and Schäfer (2010) thus argue for the nominals that obey the external argument generalization, a syntactic analysis is built on the that if the nominal denotes the external argument of the verb, then the layer that is responsible for introducing this layer should be present in the nominalization. These are sub-divided into episodic ones, which always project AS, and dispositional ones, which may leave these objects unexpressed:

(22) [nP *-er* [VoiceP [vP [RootP>]

On this view, all external argument *-ers* (agents, holders, experiencers, ...) involve (22). In (22), the *n*-layer in (22) is clearly the nominalizer. The main function of this head is to introduce the *R*-argument and in this particular case is spelt out as *-er*.

(23) teach (*x* (*y*)) teacher (*R* = *x*) such that *x* teaches *y*

Since all *-er* nouns are referential, *R* is introduced in *n*, irrespectively of the [$\pm$ event] classification.

This analysis is built upon the so called Voice Hypothesis (Kratzer 1996), according to which the external argument is not introduced by the verb itself, but by a semi-functional Voice-projection on top of *vP*. As mentioned above, the individual denoted by the *-er* nominal is, in its productive use, the one that is the external argument of the event entailed by it (see van Hout and Roeper 1998, Baker and Vinokurova 2009, who argue that *-er* is the external argument, cf. Ntelitheos 2012). Alexiadou and Schäfer (2010) proposed therefore that in these kinds of *-er* nominals the referential argument <*R*> binds a variable <*x*> located in Spec, Voice; this derives the “external argument generalization” and ensures the correct theta role for the *-er* nominal.

This analysis also captures the forms that are classified as middles, by for example Rappaport Hovav and Levin. Those nouns that do not obey the external argument are argued to be root-derived, for example *-er* nominals which are derived from adjectival stems (*foreigner*), prepositional stems (*upper*), denominal stems (*porker*), or measure words (*fiver*). This explains the low productivity of these forms. Concerning locative *-ers*, Alexiadou and Schäfer claim that actually what is nominalized is a covert location included in the meaning of, for example, *dine*.

But still, why would non-subject oriented nominals also involve *-er*? Arguably studies stressing the relevance of conceptual, cognitive, and pragmatic-semantic factors have a lot to contribute (cf. Ryder 1999). Booij and Lieber (2004) use the term of pragmatic pressure to explain what is happening. By “pragmatic pressure” they mean a situation in which context forces speakers to create a word but the language does not have a specific derivational means for doing so. When such pressure exists, one of two things happens, so the authors claim: either a formally more complex process (e.g. conversion or substantivization of a participle as in the Dutch and Greek counterparts of *-ee* nominals; see the discussion in the next section) is employed, which implies a higher degree of morphological complexity, or, more interestingly, the semantically closest productive affix is put to use (as in English) even if it requires a violation of the co-indexation criterion introduced in (19).

In conclusion, across languages, agent, instrument, and locative nominalizations bear the same form. For argument structure based approaches of the type illustrated here, agent and instrument cluster together, both being able to denote external arguments. Location can also be considered as related to the spatio-temporal argument some verbs have (Davidson 1967). For syntactic approaches, a particular affix is underspecified as to the locus of its insertion. For approaches based on lexical semantics, the clustering is a

result of the interpretation these affixes bear via the coindexation criterion (19) and how affixes interact with the bases they attach in terms of features.

For cognitive-oriented perspectives, however, this clustering has a different source. Ryder (1999) suggests that two conditions are responsible for this, namely salience and identifiability. Salience refers to the degree to which something is noticeable in comparison to its environment. Identifiability refers to the extent to which a participant is identifiable by mention of the event alone. Ryder, building on Langacker (1991), proposes the following saliency scale:

(24) Saliency: Agent > Patient > Instrument > Other cases

Ryder suggests that agent and instrument (Rappaport Hovav and Levin's intermediary instruments) are more likely to surface bearing *-er*, as they are clearly both identifiable in their own right, that is, the event can be construed as having an instrument as the head of the causal chain. In addition, both are also salient. Ryder also notes that in the history of English, the *-er* affix was originally limited to agents, and later it expanded to instrument referents during Late Middle English, and then to some Locations in Early Modern English. These two conditions and their interaction with event schemata denoted by the bases to which *-er* attaches help us understand the locative uses of *-er* as well.

Rainer (2005b) offers a diachronic account. According to Rainer, the formal identity of agentive, instrumental, and locative affixes has several sources, notably re-interpretation and approximation, both based on semantic shift, and instances of non-semantic motivation such as ellipsis, homonymization, and borrowing. Re-interpretation includes three stages: at first, there are only agentive formations, then some of them acquire an instrumental interpretation due to semantic shift, and finally the instrumental formations are re-interpreted as an independent word-formation process. Approximation skips the second stage.

14.5 *-EE* NOMINALIZATIONS

A similar picture to the one found with *-er* nominalizations has also been established for *-ee* nominals. The affix *-ee* in English has a variety of meanings as well. Most often, however, it creates object oriented nouns:

(25)	<i>verb</i>	<i>theta role</i>	<i>derived noun</i>
	employ	patient/theme	employee
	nominate		nominee
	address	goal	addressee

Nevertheless, there are subject oriented *-ee* nominals:

(26)	<i>verb</i>	<i>heta role</i>	<i>derived noun</i>
	escape	agent	escapee
	stand		standee

Barker (1998) further cites examples where the referent does not correspond to any argument of the noun, for example *amputee* refers to the person whose limb has been amputated. Booij and Lieber (2004) note that in other Germanic languages, for example Dutch, there is no corresponding affix to English *-ee*. The closest languages like Dutch come to form the counterpart of an *-ee* noun is by substantivizing past participles by means of *-e* suffixation. In Greek, certain nouns that correspond to agentive *-ee* nouns in English are built on the basis of the agentive affix used for the Greek counterparts of *-er* nominals. Others involve nominalized forms of participles as in Dutch:

(27)	drapetevo	drape-TIS
	escape	escapee-MASC.
	akrotiriazō	akrotiriasmenos
	amputate	amputated-MASC.

The most comprehensive analysis of English *-ee* nominals is Barker (1998). Barker argues that contrary to *-er*, an AS analysis of *-ee* is not adequate. On the basis of examples such as (24) and (23), one cannot argue that *-ee* binds the internal argument of the base verb. Barker puts forth instead a semantic analysis of *-ee*, according to which *-ee* binds an argument of the base verb under three conditions: (i) the argument is episodically linked to the verb, that is, the argument is a participant in the event denoted by the verb, (ii) it must denote something sentient, and (iii) it must lack volitionality. For the canonical uses, such as *employee*, the affix binds the patient argument instead of the agent argument, as this argument is both sentient and nonvolitional. Cases such as *standee*, *escapee*, and *amputee* are more complex. For *standee*, Barker argues that the argument is episodically linked, and sentient, and at least nonvolitional enough, so it can be subsumed under his general analysis. For *escapee*, Barker argues that the overall situation of an escape lacks a complete sense of control. Finally, for *amputee* he points out that the word describes the possessor of a limb that has been removed. The object argument entails a possessor that is both sentient and nonvolitional, hence the *-ee* form can be linked to the possessor.

According to Booij and Lieber (2004), Barker's analysis makes an excellent case that the analysis of *-ee* must take place at the level of lexical semantics. Their take on that is to adopt the framework of Lieber (2004), and propose the following syntactic subcategorization and skeleton for *-ee*:

- (28) *-ee*
 syntactic subcategorization: attaches to V, N
 skeleton: [+material, dynamic ([_{sentient, nonvolitional}], <base>)]

Crucially, from their perspective *-ee*, unlike *-er*, places two requirements on its coindexed argument. It places a strict requirement on the sentience of its coindexed argument and a weak requirement on the nonvolitionality of this coindexed argument. Consider the following derivations, taken from Booij and Lieber (2004): The noun *employee* receives the semantic structure in (29):

- (29) *employee*
 [+ material, dynamic ([_{sentient; nonvolitional-i}], [+ dynamic ([], [_i]))]
-ee *employ*

Assuming the verb *employ* is an activity verb, it has the skeletal feature [+ dynamic] and two arguments, the first of which is volitional, and therefore incompatible with the R-argument of the affix. The second argument is sentient but not necessarily volitional, and it is therefore more consistent with the semantic requirements of the affixal arguments. They are coindexed, and the R-argument then shares the “patient” reading of the coindexed base argument. Turning to the interpretation of *amputee*, the authors assume that the composed skeleton of *amputee* is the one in (30):

- (30) *amputee*
 [+ material, dynamic ([_{sentient; nonvolitional}], [+ dynamic ([], []))]
-ee *amputate*

Assuming that *amputate* is an activity verb whose first argument is sentient but volitional, and whose second argument is nonsentient, there is no good match for the semantic requirements of the affixal argument. But normally, the authors state, the second argument position of the verb *amputate* is occupied by a noun like *leg* or *arm*, which has its own two arguments, the second of which is its possessor, an argument which can be sentient and nonvolitional. In this system, semantic interpretation above the lexical level involves the successive composition and integration of skeletons, the R-argument of the affix will eventually come to an argument which is compatible with its semantic requirements, namely the possessor of the limb. And that is what ultimately gets coindexed with the affixal argument.

Heyvaert (2006) states that what appears to unite *-ee* nouns is that they establish a relationship between an entity and a process, a relationship which is comparable to that at clause level between an entity and a verb in the form of a past participle. Importantly, the various meanings that can be realized by a past participle (passive, present perfect, stative passive) help us understand why some languages in the absence of a dedicated affix use substantivization of a past participle to derive the counterparts of *-ee*, for example Dutch and Greek. What at first sight constitutes the most tricky subtype of *-ee* derivation—that which profiles an agent (e.g. *escapee*)—appears to have central aspects of the past-participial semantics, as described in Langacker (1991: 200–7, 221–5), in common

with the prototypical core of *-ee*. “Agentive *-ee* nominalizations profile the resultant state which an agentive entity finds itself in after some change and are thus ‘downstream’ with respect to the flow of time; non-agentive *-ee* nominalizations profile a terminal participant or an entity which is downstream with respect to the flow of energy. This entity may in addition be downstream with respect to the flow of time and portrayed as stative (as in *adoptee*)”.

Syntactic analyses of *-ee* nominals have been proposed in van Hout and Roeper (1998) and Marantz (1999). van Hout and Roeper argue that *-ee* takes a VP as its complement on the basis of the observation that such nouns do not surface with arguments,⁵ and neither do they tolerate adverbials. Since they do not provide argument positions, they lack functional projections such as Voice, where the external argument is introduced, and Aspect, where, according to their theory Case is licensed. However, *-ee* nominals contain a VP node:

- (31) a. *an employee by Mary
b. *a trainee with great effort

Marantz (1999) argues that *-ee* nominals are root-derived. In fact he takes the “truncation” observed in (32) as the result of root derivations:

- (32) a. nomin-ate, nomin-ee (cf. nomin-al), -ate for little v, -ee for little n
b. evacu-ate, evacu-ee, -ate for little v, -ee for little n

The observation is that *-ee* does not attach outside of affixes that verbalize the categorial root, for example, in *nomin-ee*, *-ee* attaches directly to *nomin-* and not *nomin-ate-*, which contains the verbalizing affix *-ate*. He further notes that we predict “truncation” for *-ee* given the semantics of *-ee* suffixation. The semantics of root affixation should go along with the morphophonology of affixation to the morphophonological root

- (33) a. nomin-at-or, evacu-at-or, *nomin-at-ee, *evacu-at-ee
b. *nomin-er, *evacu-er⁶

⁵ However, one does find examples such as the following, provided by Rochelle Lieber: from *Style* 2002: “The ‘gentle friend,’ however, seems to disappear from the poem at its conclusion, frozen out of the scene as the speaker turns definitively toward her disembodied lover, the addressee of her final series of speech acts: ‘shall not I, too, be, / My spirit-love! upborne to dwell with thee?’”

⁶ Note, however, that forms such as *narratee* and *enunciatee* can be found, as observed by Rochelle Lieber.

14.6 EVENT AND RESULT NOMINALIZATIONS

In this section, I will briefly turn to deverbal nominals such as *building* and *translation* that can refer both to the action of the base verb or its result. As Grimshaw (1990) pointed out, the two do have a different syntax, as result nominals lack AS, which event ones are accompanied by. I will not discuss this issue here any further, see the overview in Alexiadou (2001, 2010a, b), Borer (2013) for details.

As has been pointed in the literature (Asher 1993, Pustejovsky 1995), result nominals may denote a physical concrete object, such as *the construction is standing on the next street*, or the result-state of an action, as in *the obstruction may be temporary or permanent*. The question that arises then is whether we can predict which verb will give rise to which result interpretation and whether there are verbs that can form nominals with both interpretations.

Bisetto and Melloni (2007), and Ježek and Melloni (2011) identify different classes of verbs that yield ambiguous event/result nouns. On the one hand, we have verbs that express events that put a new entity into existence such as *create* and *construct*. On the other hand, we have verbs that can express a result state such as *isolate* and *obstruct*. The former class of verbs form nominals that can have an event and a result object interpretation only, while the latter can have nominal forms that are three ways ambiguous, that is, they refer to events, result states, and result objects. This is illustrated in (34) with Italian examples, from Ježek and Melloni (2011):

ostruzione ‘obstruction’

(34) (EVENT)

- a. Per evitare l'ostruzione del tubo i tubi stessi devono essere lavati.
‘To prevent the obstruction of the pipes, pipes must be cleaned’

(STATE)

- b. L'ostruzione può essere temporanea o permanente.
‘The obstruction may be temporary or permanent’

(RESULT-OBJECT)

- c. Questo test permette di capire esattamente dove si trova l'ostruzione.
‘This test allows to understand exactly where the obstruction is’

Nouns such as *construction* or *translation* cannot refer to the state of being *constructed* or *translated*, nor can they denote the state of existence of the construction and translation respectively, they can only refer to the physical or abstract object that is ‘created’ by the action. Ježek and Melloni, building on Rappaport Hovav and Levin (1998), argue that this relates to the fact that for verbs of creation the causing process (E1) overlaps the state subevent (E2), there is no independent access to the BECOME subevent and to the resulting STATE either. Such inaccessibility to the state—they argue—is inherited by

the nominal, which is therefore incapable of yielding a result state interpretation. On the contrary, the result state interpretation is available to those nominals which are derived from causatives implying no temporal overlap and in which a certain (reversible/transitory) state is independently represented in the temporal ordering of the event, like in *isolate*.

In more recent work, Rappaport Hovav and Levin (2010) propose that there is something special about verbs of creation and incremental theme verbs in general. Specifically, they argue that these verbs do not lexicalize scalar changes and in terms of event structure they are similar to manner verbs. Hence they should be associated with a simple event structure, that is the lexical semantic representation of these verbs does not contain a state component. If this is the case, then we expect nominals derived from verbs that lack a result state as part of their lexical meaning to not be able to refer to a result state.

The second issue relates to the way the result (state or object) reading can be achieved. Most authors take the event reading to be salient, and derive the result interpretation via a metonymic shift (see the discussion in Bisetto and Melloni 2007). For Pustejovsky (1995), nominals displaying the event/result meaning contrast are classified as complex types. That is, it is assumed that the event/result senses of nominals are an instance of lexically specified (or inherent) polysemy, an ambiguity available by virtue of the semantics inherent in the noun itself. In terms of lexical semantics, Bisetto and Melloni (2007) argue that there are two types of affixes, subject to different coindexation requirements, which are involved in the formation of event as opposed to result nominals: those that build event nouns are [–material] and [dynamic], while those that form result nominals are [± material] and necessarily involve coindexation of the internal argument and R-argument of the affix (result object interpretation) or the incremental result and the R-argument (result state interpretation).

Syntactic approaches to this phenomenon would have to assume that the result state interpretation involves nominalization of a VP, as it is nominalization of an event that leads to a state (Alexiadou 2009), while the object interpretation involves a nominalization of a root, as there is no event involved (Alexiadou 2001), see also Borer (2013). The event/AS reading is of course derived from a full verbal structure, see also Borsley and Kornfilt (2000), van Hout and Roeper (1998), Ntelitheos (2012). From this perspective, these affixes are underspecified as to their locus of insertion as well.

14.7 DE-ADJECTIVAL NOMINALIZATIONS

Deadjectival nominalizations constitute an under-studied domain of word formation. Recently, however, several authors have proposed that such nominalizations can be divided into two distinct groups: those that refer to the state an individual may be

in, S-nominalizations (e.g. *sadness*, *perplexity*), and those that refer to a quality an individual may possess, Q-nominalizations (e.g. *wisdom*, *beauty*). As Roy (2010) observes, de-adjectival nominalizations are ambiguous between the two readings, as exemplified in (35) for French. Roy further convincingly shows that S-nominals behave like AS nominals in Grimshaw's (1990) sense. For example, S-nominals, but not Q-nominals, can be modified by adjectives such as *constant*, and they obligatorily require the presence of a holder argument:

- (35) a. La popularité de ses chansons m'impressionne.
 the popularity of his songs me.impresses
 'The popularity of his songs impresses me'
 b. La popularité est une qualité qui lui fait défaut
 the popularity is a quality that to.him does default
 'Popularity is a quality that he is lacking'

Across languages, it is often the same affix that is involved in both S and Q readings.

(36) Suffixes

- a. French: e.g. *-ité*
 b. German: e.g. *-ität*, *-heit*, *-keit*, *-e*
 c. Romanian: e.g. *-ătate/-itate/-utate*, *-ețe*, *-ie*
 bunăătate (kindness), *frumusețe* (beauty), *voioșie* (joyfulness)
 d. Greek: e.g. *-sini*, *-otita*, *-ia*
 kalosini (goodness), *hideotita* (vulgarity), *omorfia* (beauty)

While lexicalist approaches to the type of polysemy found in (35) focus on the properties of the stem involved, syntactic approaches to this problem, such as the one advocated in Roy (2010), assume that the structure of Q-nominals differ from that of S-nominals in that the former lack an overtly realized external argument.

Note, however, that if a language has more than one affix to form de-adjectival nominalizations, sometimes a semantic selection effect can be observed. For instance, Alexiadou and Martin (2012) found certain correlations between suffixes and semantic content. Concerning the four French de-adjectival suffixes they studied in detail (*-erie*, *-isme*, *-ité*, *-itude*), the following generalizations can be drawn: (1) the suffix *-ité* is the unmarked suffix and can form Ns with any kind of aspectual interpretation; (2) *-erie* imposes a preference for the eventive reading; (3) *-isme* tends to force the de-adjectival noun to have a quality (or dispositional) reading; (4) *-itude* forces the noun to denote habits or attitudes and thereby imposes the feature of animacy and the individual-level reading. Similar observations hold for other Romance languages. The striking observation made is that de-adjectival nominalizations of type (2) have an eventive interpretation in the absence of an verbal stem. This led Alexiadou and Martin to propose that *-erie* can be decomposed into *-er*, signalling verbal word formation out of an adjective, and *-ie* signalling nominal derivation.

For English, it is generally assumed that *-ness* is the most productive affix to form de-adjectival nominalizations, while *-ity* is less productive and often gives rise to idiomatic meanings, see Aronoff (1976). For this reason, Marantz (2001) proposed that *-ity* is an affix that attaches to the root, while *-ness* is an affix that attaches to an adjective.

CHAPTER 15

VERBAL DERIVATION

ANDREW KOONTZ-GARBODEN

THE issues in the study of verbal derivational morphology are numerous and have been at the forefront of work in linguistic theory since the early days of generative grammar. In this chapter, I give an overview of some of these issues, and then focus in more detail on one particular issue that is much debated in current research—the syntactic and semantic representation of verbal derivational morphology.

15.1 THEMES IN THE STUDY OF VERBAL DERIVATIONAL MORPHOLOGY

15.1.1 Polysemy and Derivation

An important issue in the study of derivational morphology generally is the treatment of polysemy, the fact that a single derivative often has multiple meanings, which often seem relatable to one another only vaguely. The problem is exemplified by recent work by Plag (1999) and Lieber (1998, 2004) in the context of the English derivational suffix *-ize*, as exemplified by the data in (1).

- (1) The various meanings of English *-ize* (Lieber 2004: 77)
- | | |
|--|---|
| make x; cause become x | <i>standardize, velarize, crystallize, unionize</i> |
| make x go to/in/on something | <i>apologize, texturize</i> |
| make something go to/in/on x | <i>hospitalize, containerize</i> |
| do/act/make in the manner of or like x | <i>Boswellize, despotize</i> |
| do x | <i>philosophize, theorize, economize</i> |
| become x | <i>oxidize, aerosolize</i> |

As Lieber discusses, and as is shown by (1), English verbs in *-ize* have a range of meaning types that are loosely connected to one another. The challenge for derivational morphology is to understand whether the fact that all of these different uses are marked by a single suffix is an accident or not.¹ Either there are several suffixes *-ize*, with different semantics, or a single one with a uniform, if largely underspecified lexical semantics (Lieber 1998, 2004, Plag 1999). The problem is a general one, and is at the heart of the typological literature on multifunctionality and semantic maps (Croft 2001, Haspelmath 2003, Croft and Poole 2008), which covers a broad range of such phenomena in cross-linguistic perspective (see, e.g., the collection of papers in the journal *Linguistic Discovery* 8.1 (2010)).

15.1.2 Conversion

Similar questions of polysemy, but concomitant with questions about morphological form, arise in cases of conversion, sometimes also called transposition (Marchand 1969, Beard 1995, Spencer 2005, in press). These are cases where a word has one meaning when used as a word of one lexical category and another related meaning when used with a different lexical category, but with no overt morphological exponent of the derivation. Examples of this phenomenon are myriad in the literature on languages claimed to lack the full range of lexical category distinctions (see, e.g., Hengeveld 1992, Bhat 1994, Jelinek and Demers 1994, Wetzter 1996, Broschart 1997, Croft 2001, Beck 2002, Enfield 2004, Evans and Osada 2005, Koontz-Garboden 2007, Kaufman 2009, Koch and Matthewson 2009; see Chung 2012 and Koontz-Garboden 2012b for recent overview discussion). It is also found in English, however, with one of the most discussed cases being that of denominal verb formation (Clark and Clark 1979, Kiparsky 1997, Plag 1999: 219–26, Hale and Keyser 2002, Arad 2003, Lieber 2004: 89–95, Harley 2005, among many others), wherein verbs are productively derived from nouns with no morphological exponent, as with the novel formations in (2) due to Clark and Clark (1979: 767).

- (2) *to porch a newspaper, to Houdini one's way out of a closet, to enfant terrible gracefully, to houseguest with Bill Dodge, to wrist the ball over the net, etc.*

The key issues in this literature have been at once the formal morphological nature of the derivation and the semantic nature of it. Morphologically speaking, the question is whether there is a phonologically null affix deriving verb from noun (Marchand 1969: ch. 5, Hale and Keyser 2002: ch. 3), whether the morphological relationship is best characterized in some other way (Clark and Clark 1979, Lieber 2004), or whether the

¹ One of the best ways to know, which so far as I can tell has not been pursued in the literature on English *-ize* is to look cross-linguistically at other verb-forming suffixes to see if the same polysemy recurs. If it does, it almost certainly cannot be an accident (see, e.g., Haiman 1974, Haspelmath 2003, Koontz-Garboden forthcoming for relevant discussion). A similar constellation of meanings is found, for example, with the *-pa-* and *-ta-* verbalizing morphology in the Misumalpan language Ulwa (Koontz-Garboden 2009b).

superficial morphological similarity actually masks an underlying heterogeneity in the nature of the relationships (Kiparsky 1982a, 1997). This issue is tied up with the characterization of the semantic relationship—whether there are restrictions on it, and what this says about the morphological nature of the derivation, as discussed in particular detail by Kiparsky (1997).

15.1.3 The Nature of Affixal and Root Meaning

As has been noted repeatedly (Carstairs-McCarthy 1992, Rappaport Hovav and Levin 1998, Lieber 2004), a barrier to the semantic study of word formation has been development in the understanding of lexical semantics in general. Recent years, however, have seen major development in this area, with landmark studies by Dowty (1979, 1991), Hale and Keyser (1987), Pinker (1989), Jackendoff (1990), Levin (1993), Levin and Rappaport Hovav (1995), Van Valin and LaPolla (1997), and Rothstein (2004), among others.

As discussed further in Section 15.2, what unifies many approaches to the study of lexical meaning is some kind of semantic decomposition, much in the tradition of generative semantics (Lakoff 1965, McCawley 1968), wherein the meanings of words are decomposed into primitives that are either unique to that word, often called roots (Rappaport Hovav and Levin 2010), constants (Grimshaw 1993), or the body (Lieber 2004) and those that are shared with other words, varyingly called the template (Rappaport Hovav and Levin 1998), semantic structure (Grimshaw 1993), or the skeleton (Lieber 2004). Regardless of nomenclature, the basic idea is the same—what unites classes of lexemes, such as the verb classes of Levin (1993), is broad, shared elements of meaning, whereas what differentiates members of the class from one another are different root/constant elements of meaning. Classes of lexemes, for example, change of state verbs (e.g. *break*, *split*, *crack*) and manner verbs (e.g. *run*, *swim*, *crawl*) in Rappaport Hovav and Levin (2010), differ fundamentally from one another in having different templatic/structure type meaning.²

Assuming that there are generalizations about the kind of meaning that a particular affix introduces, such meaning is by definition the kind of meaning that generalizes across lexemes, given that affixes appear with multiple lexemes, and will therefore be templatic in nature.³ It will contrast with the meanings of individual morphological roots, which although they may carry templatic meaning, always have their own

² See Section 15.2 for further discussion of the root/template distinction more generally.

³ This holds only in those instances where a derivational operation introduces new lexical entailments. It is likely that there are at least some derivational operations that do not do this, but rather that simply effect the saturation and quantification of some variable. Deverbal adjective formation, for example, can be viewed as such an operation (see Koontz-Garboden 2010 and references there). In the verbal domain, what Dowty (1979) calls detransitivization (also called unspecified object deletion and

idiosyncratic element of meaning that distinguishes them from other lexemes in a class. Lieber (2004) develops precisely such a view.

15.1.4 Inflection, Derivation, and Lexicalism

Although derivational morphology is traditionally contrasted with inflectional morphology, the criteria for distinguishing one from the other are notoriously problematic (Anderson 1982, Stump 1998). From the perspective of the syntax and semantics of verbal derivational morphology and lexicalist theories building on Chomsky (1970) (Chomsky 1981, Pollard and Sag 1994, Bresnan 2001), whether a particular morpheme is categorized as derivational or inflectional is not particularly important, as the strong lexicalist hypothesis has it that the internal structure of words is not visible to syntactic operations at all (Anderson 1982: 573). In the 1980s this position gradually weakened in the Chomskyan literature, and there developed a trend toward syntactification of inflectional morphology (Anderson 1982: 587), particularly in the form of functional heads, best represented perhaps by the work of Pollock (1989). Gradually, especially thanks to the work of Baker (1985b, 1988), more and more morphology became syntactified, in particular grammatical function changing morphology like causative (Baker 1985b, 1988), applicative (Baker 1988), and passive (Baker et al. 1989).⁴ Although such proposals have given rise to much controversy both for (Alsina 1992, Sells 1995, Bresnan and Mchombo 1995, Bresnan 1996, 2001) and against (Marantz 1997, Pytkänen 2002, Embick 2004) lexicalist positions, there is now a robust and influential corner of the non-lexicalist literature in which it is generally taken for granted that most if not all morphology is syntactically represented. Examination of this issue, particularly in the domain of derivational phenomena, where it is more recent and more controversial, is at the forefront of much current work on the syntax and semantics of derivational morphology, particular as it relates to verbs. In the remainder of this chapter, I examine this issue in detail, drawing on a particular case study in doing so.

15.2 TWO WAYS OF APPROACHING THE SYNTAX AND SEMANTICS OF DERIVATIONAL MORPHOLOGY

One of the central issues in the study of verbal derivational morphology since the days of Generative Semantics, and one which has re-emerged in recent years, is the issue of

indefinite object drop, e.g. *Kim ate cake* vs. *Kim ate*) is such an operation, to the extent it is viewed as a derivational operation (in spite of the lack of overt morphology), as it is by Dowty (1979: 308).

⁴ See Belletti (2003) and Roberts (2003) for further overview and discussion.

lexical decomposition and what the relationship is between verbal derivational morphology and compositional structure. The program of lexical decomposition (for relatively recent discussions see, e.g., Hale and Keyser 1987, Pinker 1989, Jackendoff 1990, Grimshaw 1993, Levin 1993, Pesetsky 1995, Van Valin and LaPolla 1997, Rappaport Hovav and Levin 1998, Lieber 2004, Harley 2005, Levinson 2007, Ramchand 2008, Beavers and Francez 2012, among many others), as briefly discussed in Section 15.1.3, is about understanding what elements of word meaning (lexical entailments as in Dowty 1989, 1991) the syntax and morphology of languages are sensitive to. To take an example from Grimshaw (2005: 75–6), consider the two verbs in the sentences in (3).⁵

- (3) a. The ice cream melted.
b. The ice cream froze.

There are certain aspects of the meanings of each of the verbs in (3) that are shared and which grammatical processes are sensitive to, and others which are not. As Grimshaw says

that melt and freeze both [can] mean to change state is linguistic, that they concern changes in liquidity, and that each means what it means and not what the other means is not... [T]he aspect of meaning that distinguishes... melt from freeze is of no linguistic significance and plays no role in the grammatical system of the language. (Grimshaw 2005: 76)

The kinds of process that Grimshaw is referring to include grammatical alternations like the causative alternation and others (see Levin 1993) and morphological marking (e.g. English *-en* which derives change of state verbs from adjectives as in *red*den from *red*). The key point is that the kinds of meanings that a verb has are of two kinds. The first is the kind that grammatical processes are sensitive to and which classes of verbs like those identified by Levin (1993) are identified on the basis of. I call this “templatic” meaning following Levin and Rappaport Hovav (2003). The second kind of lexical meaning component is idiosyncratic and often called “root” meaning, as discussed in Section 15.1.3. These names are mnemonic for the nature of their representation in a decompositional structure, with a verb like *melt* in (3a) having a decomposition like (4), which factors out the change of state meaning.⁶ This leaves all of the other meaning components, principally the stative core, packaged in the root *mel*ted, which is the core lexical semantic unit which distinguishes a verb like *melt* from a verb like *freeze*.

⁵ I have changed Grimshaw’s examples from causative to inchoative verbs for rhetorical purposes. This also simplifies the illustration of the point she makes in the quote that follows.

⁶ In order to keep the representations from being more complex than is necessary for the purposes of this discussion, I give a decomposition like the ones in Levin and Rappaport Hovav (1995) and Rappaport Hovav and Levin (1998), which gloss over the functional nature of verbs and obscures their role in compositional semantics. See Dowty (1979) for compositionally more realistic lexical decompositions.

(4) [BECOME melted(y)]

A key issue in this literature is how exactly templatic meaning is represented and compositionally contributed to the meaning of a sentence. There are various intermediate positions, but at the extremes, lexicalists have it that lexical decompositions can be manipulated only in the lexicon, and that the decompositional meaning is not syntactically represented.⁷ Early work in Government-Binding theory (Chomsky 1981) and work in Head-Driven Phrase Structure Grammar (Pollard and Sag 1994), Lexical-Functional Grammar (Bresnan 1982), and Lexical Decomposition Grammar (Wunderlich 1997, Kiparsky 2001) all fall into this category, though there is much more than it is possible to cite here as well. Various non-lexicalist approaches to the syntax/semantics interface have it, by contrast, that all templatic meaning is syntactically represented, and it is only the content meaning that is lexically represented (Borer 2005a, b, Embick 2009).⁸ This position is stated particularly clearly by Embick (2009), as a general working hypothesis in the non-lexicalist tradition that he calls “the bifurcation thesis for roots”:

If a component of meaning is introduced by a semantic rule that applies to elements in combination, then that component of meaning cannot be part of the meaning of a root. (Embick 2009: 1)

Stated in the terms laid out above, the idea is that all decompositional meaning is introduced syntactically by functional projections. Only idiosyncratic meaning that does not figure into grammatical processes is packaged into the morphological root, itself a syntactically represented object on such theories. One of the key differences between the two positions, then, is in whether templatic meaning is taken to be syntactically or lexically represented.

Derivational morphology is often taken as evidence for a particular unit of templatic meaning in a morphologically derived word, whether lexically or syntactically represented (see, e.g., Levin and Rappaport Hovav 1998). For example, the *-en* morphology in English is commonly assumed to introduce change of state meaning, represented compositionally as the BECOME operator (and its cross-theory kin) (see, e.g., Dowty 1979: 307; Embick 2004: 365, among many others). On the surface at least, it appears that it takes a state denoting base like those in (5a) and returns a change of state denoting verb, like those in (5b), hence the common naming of the latter as “deadjectival.”

⁷ See Chomsky (1970) for the origins of this idea and Dowty (1979) for particularly lucid critical discussion.

⁸ Midway positions are, of course, both conceivable and attested in the literature; Alexiadou et al. 2006, for example, propose classes of roots with certain selectional properties. Baker (2003), while otherwise generally adhering to syntactified decomposition, at least in the verbal domain that concerns us here, also allows for some lexical derivation, noting that “once the syntactically predictable morphology has been stripped away, there remains a residue of morphology that seems to have nothing to do with syntax” (Baker 2003: 280). He therefore proposes that “what is inserted into an X⁰ slot can be a root, a derived stem, or an inflected word” (Baker 2003: 289).

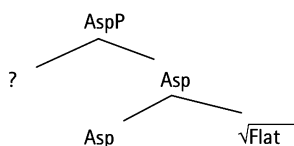
- (5) a. *awake, bright, broad, cheap, coarse, damp, dark, deep,...*
 b. *awaken, brighten, broaden, cheapen, coarsen, dampen, darken, deepen,...*

The status of morphologically unmarked lexemes that derived ones are morphologically related to, as with (5a, b), is a key point of difference between the two kinds of theory when it comes to derivational morphology, particularly in relation to verbs. There are two classes of unmarked lexeme that provoke controversy. The first is the unmarked lexeme that looks, on the surface, like it is the input to derivational operations. The adjectives in (5a) fall into this class. The second class is composed of morphologically simple lexemes that share the templatic meaning of derived lexemes. Morphologically simple change of state verbs like *melt* and *freeze* in (3) fall into this class in relation to the morphologically derived change of state verbs like those in (5b).

In a typical lexicalist approach (e.g. Koontz-Garboden 2006), the unmarked lexeme is taken as lexically listed, even if its meaning (as it often does) includes templatic entailments, and the derivational morphology is taken to operate on the underived form to yield the derived form. This is the case not only morphologically, but also semantically. That is, on this view, the surface morphology reflects semantic composition directly, so that the meaning of the derived form is a function of the meaning of the surface underived form, the surface morphology, and the way in which the two are put together. What this means for the relationship between adjectives and deadjectival change of state verbs illustrated in (5) is that the verbs, both formally and semantically, are derived from the adjectives. That is, *red*den is derived from the adjective *red* by way of an operation that results in the suffixation of *-en* morphology and the addition of change of state entailments (in the form of the decompositional BECOME operator). What this view entails for *melt* and *freeze* when compared to the deadjectival change of state verbs in (5b) is that while the latter have their change of state meaning in virtue of a derivational operation, the former have their change of state meaning lexically; they are simply change of state verbs in the lexicon, consistent with their morphologically simple form. Theories that adhere to this kind of view, then, take the surface morphology to reflect underlying semantic derivation transparently. Because of this, I call such views in what follows WYSIWYG approaches (for ‘what you see is what you get’).

In non-lexicalist approaches that adhere to something like the bifurcation thesis for roots, by contrast, the unmarked lexeme, if its meaning includes any templatic entailments, will still be treated as derived from a more basic unit (the root; Arad 2003, Harley 2005, 2011, Levinson 2007). Even in cases where the unmarked lexeme actually has no templatic meaning, as is the case for the adjectives in (5a), on theories that adhere to principles of Distributed Morphology (Halle and Marantz 1993, Marantz 1997), the unmarked lexeme will still be derivationally complex, given the leading idea that all words are formed syntactically from morphologically bound pre-categorial roots (see Arad 2005: ch. 1 for an overview). In the case of adjectives like (5a), for example, Embick (2004) argues that syntactically, they are as in (6), derived from state-denoting bound precategorial roots that merge with an adjectivizing phonologically null functional head which he calls Asp.

(6) The adjective flat in DM (Embick 2004: 363; Embick's ?)



These ideas have a number of consequences for derivational morphology. First, in many cases, they entail that the morphologically derived form is not actually derived from the morphologically unmarked form, but rather that the two are related to one another in a more indirect way. To return to the example of adjectives and change of state verbs that are morphologically related to them, English adjectives are actually derived lexemes on this approach, as is clear from (6). The change of state verb is not actually derived from the adjectival structure, but rather from the root in combination with verbalizing functional structure. Because some of these functional heads are taken to be phonologically null in English, the derivational relationship is obscured between adjective and verb, but the morphological prediction is that there should be languages in which it is not, so that it can be transparently seen that “deadjectival” verbs are not really derived from adjectives, but rather from the roots that adjectives are also derived from.

A second consequence of this kind of non-lexicalist view is illustrated by the contrast in the morphological complexity of change of state verbs like those in (3) and those in (5b). The templatic meaning of these two classes of verb is identical. As already discussed, the fact that one class is morphologically complex, while the other is morphologically simple can be captured by treating the derived class as having its templatic meaning by virtue of a derivational operation, and the simple class as having the templatic meaning that it has by virtue of lexicalization. Analyses that adhere to the bifurcation thesis for roots, however, give a different treatment to these two classes. Since they have identical templatic meaning, and since all templatic meaning is introduced syntactically, both classes are treated as having their templatic meaning as a consequence of derivational operations. The differences in overt morphological complexity are, on this view, language-specific accidents. Again, there are cross-linguistic predictions that follow as a consequence of this view. The norm, for reasons I discuss in more detail below, absent any intervening external factors, is predicted to be that all verbs with the same kind of templatic meaning have the same morphological behavior cross-linguistically, since they have a uniform syntactic representation. As a consequence, on an approach like this the direction of derivation reflected overtly in the derivational morphology is not necessarily representative of the underlying semantic and syntactic direction of derivation. On at least some of these approaches, there is necessarily, by virtue of a priori assumptions regarding the nature of the syntax/semantics/morphology interface, more derivation than the surface morphology betrays. Because the notion of the root is central to these views not only in a semantic sense (as is normal in decompositional approaches), but also in a morphosyntactic sense, I call such approaches in what follows “root-based approaches.”

As Harley (2012: 2) notes, the syntactic, morphological, and semantic derivation of change of state verbs like those in (3)—has been one of the core empirical domains for theoretical debate in relation to lexical and syntactic decomposition since the onset of Generative Semantics (see, e.g., Lakoff 1965, McCawley 1968, Fodor 1970, Dowty 1979, among others). And nowhere is the contrast between the two approaches to the syntax and semantics of verbal derivational morphology clearer. In the section that follows, I build on the discussion already laid out above and set out some of the major contrasts of these two approaches by examining change of state verbs and the contrasting morphological and semantic predictions that these two competing theories make in relation to the causative/inchoative alternation.

15.3 CAUSATIVE/INCHOATIVE ALTERNATION

The causative alternation (see Schäfer 2009 for an overview) is a verbal alternation in which the same change of state verbal lexeme can be used in both a transitive frame (7a) and in an intransitive frame (7b).

- (7) a. Kim broke the vase.
b. The vase broke.

While in English, there is no overt derivational relationship between the causative and inchoative variants in the alternation, the situation is widely acknowledged to be different cross-linguistically (Nedjalkov and Silnitsky 1973, Haspelmath 1993, Nichols et al. 2004, inter alia) where derivational relationships between the variants are overtly observed. In some cases, derivation of the causative from the inchoative is observed, as with the Tongan (Polynesian) data in (8).

- (8) a. *lahi*
 ‘become big’
b. *faka-lahi*
 ‘cause to become big’ (Koontz-Garboden 2005: 83)

In other cases derivation of the inchoative from the causative is observed, as with the Eastern Armenian data in (9).

- (9) a. *epel*
 ‘cause to become cooked’
b. *ep-v-el*
 ‘become cooked’ (Megerdooonian 2002: 98)

There are also what Haspelmath (1993: 91) calls “non-directed” derivations, where both causative and inchoative are separately derived from a root, as with the Warlpiri data in (10).

- (10) a. *wiri-jarri-*
 ‘become large’
 b. *wiri-ma-*
 ‘cause to become large’
 c. *wiri*
 ‘large’
- (Hale and Keyser 1998: 93)

Additionally, there are idiosyncratic suppletive relationships between causative and inchoative, as with the English *kill* (causative) and *die* (inchoative).

Although the situation looks chaotic, the typological and lexical semantic literatures suggest that there are generalizations to be made about direction of derivation that have to do with the nature of the event underlying the causative/inchoative pair. One of the most important observations of this body of work is that across a range of languages, there is a tendency for verbs naming certain event types to differ in the derivational processes that they undergo. When the alternation is directed, so called “externally caused change of state verbs,” verbs naming events like breakings, splittings, crackings, etc. (Smith 1970, Croft 1990, Haspelmath 1993, Levin and Rappaport Hovav 1995, McKoon and Macfarland 2000, Wright 2001) tend to have the causative in the morphologically basic form, with the inchoative derived from it (via the “anticausative” derivation). Internally caused change of state verbs, however, which name events that tend to come about on their own by virtue of the internal properties of the undergoer of the change of state event—events like blossoming, rusting, sprouting, fermenting, and other “entity specific changes” (Levin 1993: 247)—tend to be lexicalized as inchoatives, with the causative overtly derived (via the causative derivation). Famously, morphological contrasts can be observed internal to single languages, with the direction of derivation being different according to event type (see, e.g., Haspelmath 1993, Nichols et al. 2004).

For example, O’Odham and Quechua both exhibit causativization (11)–(12) and anti-causativization (13)–(14), with the direction of derivation differing according to the nature of the change of state event in question, a fact illustrated by the contrasting data below.

- (11) O’Odham causativization (Hale and Keyser 1998: 92)
- a. *weg-i-(ji)d*
 ‘cause to become red’
 b. *weg-i*
 ‘become red’
- (12) Cuzco Quechua causativization (Cusihuaman 1976: 230)
- a. *wirayay*
 ‘become fat’
 b. *wiraya-chi-y*
 ‘cause to become fat’

- (13) O’Odham anticausativization (Hale and Keyser 1998: 97)
- a. *mul*
‘cause to become broken’
 - b. *è-mul*
‘become broken’
- (14) Cuzco Quechua anticausativization (Cusihuaman 1976: 166)
- a. *wisq’ay*
‘cause become closed’
 - b. *wisq’a-ku-y*
‘become closed’

Proponents of WYSIWYG approaches take these contrasts as meaningful, and take the differences in direction of derivation to reflect underlying differences in semantic derivation. They cite in support of this claim subtle semantic and syntactic contrasts showing that morphologically derived inchoatives and underived inchoatives often contrast in their syntactic and semantic properties, in ways that suggest that the derived inchoatives share some component of meaning from the causative variant that is not necessarily encoded in many morphologically simple inchoatives (see Labelle 1992, Centineo 1995, Levin and Rappaport Hovav 1995, Alexiadou et al. 2006, Schäfer 2008, Koontz-Garboden 2009a). Koontz-Garboden (2009a: 106–10, 112–19) provides an overview of diagnostics internal to Spanish, showing that at least some underived inchoatives and derived inchoatives contrast with one another.⁹ For example, it has been claimed, that the adverbial modifier *por sí solo* ‘by itself’ can only modify verbs with causative lexical entailments (see Chierchia (2004) on Italian for the original claim). It is fine with transitive causative verbs (15a), but unacceptable with stative predicates (15b), as expected if this claim holds true.

- (15) a. *no se puede decir que ninguno de los golpes haya matado por sí solo*
no REFL can say that none of the hits has killed by self only
a la víctima
to the victim
‘It cannot be said that no hit has by itself killed the victim’ (<<http://www.gacetajuridica.com.pe/boletin-nvnet/abro5/boletin18-04.htm>>)
- b. **El carro es rojo por sí solo.*
The car is red by REFL only
*‘The car is red by itself.’ (Koontz-Garboden 2009a: 107)

⁹ See Horvath and Siloni (2011) for rebuttal and Beavers and Koontz-Garboden (in press) for further discussion and justification.

This diagnostic contrasts at least some morphologically underived inchoatives (16) with inchoatives derived with the reflexive *se* (17), suggesting that the former class lacks a causative meaning component, while the latter has one.¹⁰

- (16) a. ??*Juan empeoró por sí solo.*
 Juan worsened by REFL only
 ‘Juan worsened by himself’ (Mendikoetxea 1999: 1598)
- b. ??*La leche hirvió por sí solo.*
 The milk boiled by REFL only
 ‘The milk boiled by itself’ (Mendikoetxea 1999: 1598)
- c. ??*El niño creció por sí solo.*
 The child grew by REFL only
 ‘The child grew up by itself’ (Mendikoetxea 1999: 1598)
- (17) a. *El barco se hundió por sí solo*
 the boat REFL sank by REFL only
 ‘The boat sank by itself’ (Mendikoetxea 1999: 1594)
- b. *La puerta se abrió por sí solo*
 the door REFL opened by REFL only
 ‘The door opened by itself’ (Mendikoetxea 1999: 1593)
- c. *La ruptura continuó alrededor de esta barrera, pero treinta segundos después, cuando había avanzado 200 km, el duro bloque de la barrera se rompió por sí solo.*
 the rupture continued around of this barrier but thirty seconds after when had advanced 200 km the tough block of the barrier REFL broke by REFL self
 ‘The rupture continued around the barrier, but after thirty seconds, when it (the rupture) had advanced another 200 km, the tough block of the barrier broke by itself.’ (<<http://www.scidev.net/en/middle-east-and-north-africa/news/peru-quake-study-could-help-predict-future-disaster.html>>)

Koontz-Garboden (2009a) discusses other diagnostics that show similar contrasts, including behavior with other adverbial modifiers, and behavior in the scope of negation, which similarly suggest a contrast between derived and underived inchoatives. The data above show a language internal contrast. But they do nothing to address the question whether the cross-linguistic variation in direction of derivation is meaningful

¹⁰ See Koontz-Garboden (2009a: 109, fn. 27) and Beavers and Koontz-Garboden (in press) for discussion of the fact that not all underived inchoatives are expected to contrast, given the nature of lexicalization.

or not. For example, many of the verbs in Spanish that show an anti-causative derivation show the causative derivation in Indonesian, as shown by the data in (18). Are these morphological differences reflective of differences in the semantic direction of derivation and/or the nature of the underlying events?

- (18) a. Spanish anti-causativizing verbs
romper(se) ‘break’; *cerrar(se)* ‘close’; *ahogar(se)* ‘sink/drown’; *destruir(se)* ‘destroy’; *acabar(se)* ‘finish’
 b. Indonesian causativizing verbs (Haspelmath 1993: 116)
patah/me-matah-kan ‘break’; *tutup/me-nutup* ‘close’; *tenggelam/me-nenggelam-kan* ‘sink’; *binasa/mem-binasa-kan* ‘destroy’; *selesai/me-nyesai-kan* ‘finish’

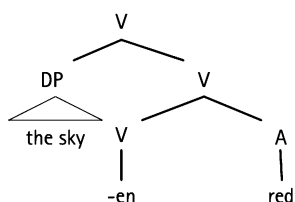
On the WYSIWYG analysis, the null hypothesis would be that the differences in (18) reflect genuine differences in the semantics of the causative alternation in these two languages. So, for example, such a view would have it that while causative *break* is semantically derived from the inchoative in Indonesian, the syntactic and semantic direction of derivation in Spanish is precisely the opposite, as reflected by the morphology. The prediction the WYSIWYG analysis makes is that although the translations into English are identical, the lexical semantics of the verbs in the two languages might differ from one another in subtle but linguistically substantive ways. Because the Spanish inchoatives in (18) are derived from causatives, the WYSIWYG approach predicts that the causative meaning present in the base form is maintained in the meaning of the derived form.¹¹ This meaning should be detectable, as illustrated above, in the form of diagnostics linked to causative meaning, such as modification by *por sí solo* ‘by itself’, behavior in the scope of negation, and others. The Indonesian ones, by contrast, would not be expected to uniformly pass these diagnostics. Since they are lexicalized as inchoatives, there is no necessary expectation that they will encode causative meaning (though nothing precludes it). Because of this, the theory allows that there could be a contrast, so that some of them would be lexicalized without causative meaning, with the causative morphology introducing this meaning derivationally. Again, this could be tested with well-motivated diagnostics for causative meaning, the prediction being that unlike the Spanish derived inchoatives, the Indonesian underived inchoatives might not uniformly pass causative diagnostics.¹² To the best of my knowledge, no cross-linguistic work has been carried out at the requisite level of semantic detail to test these predictions.

¹¹ At least assuming the widely adhered to Monotonicity Hypothesis, whereby decompositional operators are not deleted by derivational operations (Koontz-Garboden 2012a).

¹² As noted in footnote 11, given that lexicalization is idiosyncratic, absent an understanding of the semantics of the causativization process itself, there is no prediction that underived inchoatives will or will not have causative entailments. The strong prediction is only in the other direction—an inchoative derived from a causative necessarily retains the causative meaning. The prediction for underived

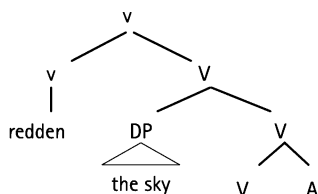
On root-based views, rather than reflecting cross-linguistic differences in semantic derivation, the observed differences are generally taken to superficially mask underlying uniformity. There are at least two such types of theory. On one, articulated particularly clearly by Hale and Keyser (2002), causative is derived from inchoative, and inchoative is derived from stative. The basic idea is that all change of state verbs are derived syntactically (and therefore semantically) from a state denoting lexeme (whether these are adjectives, bound roots, or some other lexical category). The inchoative is then derived directly from this lexeme with a verbalizing functional head. This is illustrated for English by the derivation in (19).

- (19) English inchoative *red*den (as in *The sky reddened*) (Hale and Keyser 2002: 48)



The causative, in turn, can be derived from this structure with a phonologically null causativizing functional head into which the verb moves, after first picking up the inchoativizing *-en* morphology via movement through that functional head.¹³

- (20) English causative *red*den (as in *The sunset reddened the sky*) (Hale and Keyser 2002: 48)



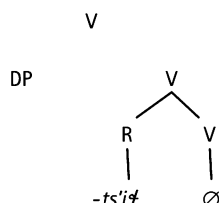
Although the morphology may differ across languages, unlike what would be presumed in the WYSIWYG view, the derivational relations are identical. What accounts for the observed differences in morphological derivation is not an actual difference in syntactic or semantic derivation, but rather differences in the realization of the verbalizing

inchoatives is much weaker—they do not necessarily have (but may have, given the idiosyncratic nature of lexicalization) causative entailments.

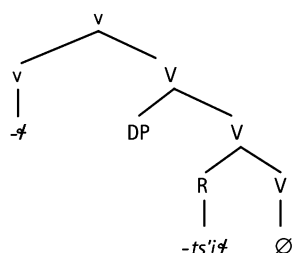
¹³ Hale and Keyser (2002: 48) actually give what I have reproduced as a “little-*v*” in (20) as a note labelled *V*. Given that elsewhere in comparable structures, e.g. (22) they give the causativizing functional head as a little-*v*, I suspect this is a typo, and have thus given it as a little-*v* in (20) below.

functional heads. So, while Hale and Keyser treat the English inchoativizing functional head as overt (*-en*) and the causativizing as null, exactly the reverse is what they assume for Navajo, as shown by their derivations, for the verb meaning ‘shatter,’ in (21) and (22) (where “R” labels the root).

- (21) Navajo *-ts’it̚* ‘shatter (inchoative)’ (Hale and Keyser 2002: 113)



- (22) Navajo *-t̚-ts'it̚* ‘shatter (causative)’ (Hale and Keyser 2002: 114)



In short, on the approach in Hale and Keyser (2002), across all change of state verbs, an inchoative structure always underlies a causative structure (see Hale and Keyser 1998 for a contrasting view). Surface differences in morphology are a consequence of differences in realization of functional heads—while some are realized as phonologically null, others are overt. In neither English nor Navajo are both the inchoative and causativizing functional heads phonologically overt. In English the former is overt, in the form of *-en*, while the latter is phonologically null. By contrast, in Navajo, it is the inchoativizing head that is phonologically null and the causativizing head that is overt. More generally, the cross-linguistic morphological prediction is that in the general case, both should be overt. There might be some language-particular accidents, as is the case on this view for English causativizing heads, and Navajo inchoativizing heads (both of which are phonologically null), but these should not repeat themselves in any systematic fashion. This is because if a putative derivational morpheme/functional head is systematically phonologically null cross-linguistically, then it calls into question its existence, since given the arbitrariness of the sign, there is no more reason to believe a particular functional head would be phonologically null than there is to believe that it would be commonly realized in any other particular way. In other words, the phonological shape of functional heads varies arbitrarily across languages. If the only cases where they do not are those where

they are null, then there is cause for suspicion about whether the null functional heads actually exist in the first place.

A second type of root-based analysis has it that both causative and inchoative are actually separately derived from a state-denoting root, as is overtly the case for at least some causative/inchoative pairs in at least some languages, viz. the Warlpiri data in (23).

(23) Warlpiri (Hale and Keyser 1998: 93)

- a. *wiri-jarri-*
'become large'
- b. *wiri-ma-*
'cause to become large'
- c. *wiri*
'large'

This is the approach pursued by Piñón (2001).¹⁴ Piñón posits a root that denotes what he calls “a causative–inchoative pair”—a pair that has a causative meaning as one member, and an inchoative meaning as the other. Causative and inchoative are each separately derived from this root, causative by picking out the causative meaning and inchoative by picking out the inchoative meaning. In this way, he explains the presence of both causativizing and anti-causativizing morphology cross-linguistically. Reversals in the direction of derivation like those in (11)–(14) are accidental on this approach, a consequence simply of the idiosyncrasy of morphology. What accounts for these is phonologically null morphology—when there appears to be a causativizing direction of derivation, the morphology deriving the inchoative from the state denoting root is simply phonologically null. When there appears to be anti-causativizing morphology, it is simply because the morphology deriving the causative from the state denoting root is phonologically null. In short, all causative/inchoative alternates on such an approach are like Warlpiri in (23), whether it overtly looks like this or not. And because of this, approaches like this make a starkly different prediction about the meaning of the inchoative when compared to the WYSIWYG approaches: because the inchoative is not derived from the causative (but rather from a more abstract root), the inchoative is not predicted to encode causative entailments. Another more purely morphological prediction of analyses like these is that the event-based morphological generalizations of Haspelmath are accidental, since the claim is that the derivation is actually uniform (both causative and inchoative derived from an underlying root; non-directed, in Haspelmath’s terms). Any deviation from this is a consequence of phonologically null

¹⁴ Piñón’s discussion is mostly about the compositional and lexical semantics; there is no discussion about how exactly he envisions his analysis at the syntax/semantics interface. As a consequence it is not really possible to classify it as either lexicalist or non-lexicalist. The point important in the present context, however, is that it is clearly organized around a notion of morphological root, and also aims for cross-linguistic uniformity, like some other root-based approaches.

morphology, which is by definition an accident in the same way that the particular phonological form realizing any meaning in any language is accidental (as discussed above). And being an accident, it is not expected to repeat itself in any systematic fashion. Therefore, if this view is correct, we should not find the absence of morphology with the same kinds of causative/inchoative pairs in case after case across languages. Some might argue that Haspelmath's (1993) observations already counter-exemplify this prediction, since what they show is exactly this—that direction of derivation is indeed conditioned by event type, as discussed above, with particular kinds of events generally showing an anti-causative derivation between causative and inchoative and other kinds of events generally showing a causative direction of derivation. Others might argue, however, as Haspelmath (1993: 96–7) himself notes, that his sample is biased toward Indo-European languages, and that as Nichols et al. (2004) argue, this group of languages has much more anti-causativization than is the norm cross-linguistically. In fact, the more balanced sample provided by Nichols et al. (2004) does provide a picture that looks more morphologically chaotic. At the same time, it does not include the diversity of event types that Haspelmath's survey does. As a consequence, it still seems an open question the extent to which direction of derivation really is conditioned by event type across languages. As is clear from the discussion above, a lot hinges on this particular issue.

In sum, at least as concerns the causative/inchoative alternation, it is very much an outstanding question in the literature what exactly the syntactic/semantic significance of derivational morphology in this domain is. While some assume that derivational morphology transparently reflects the underlying direction of semantic derivation, and that the observed cross-linguistic variation reflects genuine cross-linguistic differences in lexical semantics, others assume that there is a much less direct relationship and that cross-linguistic variation in this area does not implicate cross-linguistic differences in lexical semantics. Although there is much study of the causative alternation, there has actually been fairly little attention paid to the kinds of data that would decide the issue. The kinds of data that need to be looked at are both the lexical entailments, as described above, and the directions of derivation in relation to different event types across languages, in the spirit of Haspelmath (1993) and Nichols et al. (2004), but on a larger scale. A study with typological balance of Nichols et al. (2004) is needed, but with more change of state verbs, and of more diverse types.

15.4 DISCUSSION AND CONCLUDING REMARKS

The contrasts between root-based and WYSIWYG views of the morphology/syntax/semantics interface laid out above are sharp. Although there are certainly contrasts in semantic and morphological predictions of these two types of approach in other areas of verbal derivational morphology (notably in the derivation of states and changes of state; Koontz-Garboden 2011), they are often less stark. Change of state verbs bring

the predictions of the two approaches into such clear relief thanks to the fact that all languages have ways of expressing change of state events, and the ways in which they do this are related to one another in fairly transparent ways. Given that the syntax and the (lexical) semantics are tightly linked on root-based approaches, and given that change of state semantics seem to be universally expressible, the strongest root-based approaches have it that the syntax of change of state events is also universal, as we have seen above. Since the derivational morphology follows from the syntactic structure on these approaches, the root-based prediction, at least of the strongest root-based views examined above, is that absent language specific accidents (e.g. phonologically null morphology), the direction of derivation (with the causative alternation) should be cross-linguistically uniform. The WYSIWYG prediction, by contrast, is that robust contrasts in morphological complexity reflect underlying differences in lexical semantics. So, differences in direction of derivation in the causative alternation, for example, are predicted to reflect differences in lexical semantics. As discussed above, further investigation is needed to determine the extent to which these predictions are supported or falsified by cross-linguistic facts.

The situation is different with other derivational phenomena that are found in the verbal domain (see Haspelmath and Müller-Bardey (2004) for an overview of some of these). Consider, for example, applicative morphology, morphology that (in a range of ways—Pylkkänen 2002, Haspelmath and Müller-Bardey 2004, Peterson 2007) adds a direct argument to the verb's argument structure, as exemplified by the Ulwa example in (24), where *kang* adds an argument to the argument structure of the verb *daknaka* 'cut', as can be seen by comparing (24a) and (24b).

- (24) a. *Muih balna Karawala asang-ka kau pan isau pal-ka dak-dida.*
 person PL Karawala town-3SING in tree many very-ADJ cut-3PL.PAST
 'People cut many trees in the village of Karawala' (fieldnotes, 0405-460)
- b. *Una balna bai kaupak w-î Karawala pan-ka kang dak-dida.*
 mestizo PL far from come-SS Karawala tree-3SING APPL cut-3PL.PAST
 'The Mestizos came from far away and cut down
 Karawala's trees (on them)' (fieldnotes, 0405-460)

Although all languages presumably have ways of expressing the propositions that are expressed with applicatives in languages that have them (though this is perhaps a trickier question than one might presume; von Stechow and Matthewson 2008), the applicative construction forces a particular syntax onto the construction that has a range of unique properties (Baker 1988, Bresnan and Moshi 1990, Pylkkänen 2002, McGinnis 2008). Since the applicative alters templatic meaning, a root-based analysis might have it that it does so syntactically. Crucially, however, standard approaches to applicatives that are framed in root-based terms do not assume that the syntactic structures that underlie applicativized verbs (whether "high" or "low" ones, in the terminology of Pylkkänen 2002: ch. 2) are cross-linguistically universal. Rather, the standard position seems to be that "crosslinguistic variation in the semantic and syntactic types of applicatives, and in

their ability to combine with other applicatives, is... attributed to lexical parameters, based on the availability of various applicative heads in a given lexicon, and on their semantic and selectional properties" (McGinnis (2008: 1227) and Pylkkänen (2002) for a similar view). Of course, theories might differ on whether such variation is reduced to functional heads or not, but the point is that root-based and WYSIWYG approaches do not make the sharply diverging morphological predictions in this area that they do in relation to the causative alternation, since ultimately both kinds of view agree on the possibility that such operations (however represented) can simply be absent in particular languages.

What this all leads to, then, is contrasting views about major sources of cross-linguistic variation in derivational morphology. On the root-based views we have examined, we have seen two sources. One, evidenced in this section, is the availability of particular derivational operations (perhaps reducible to the availability of certain functional heads) in particular languages. The second, observed in the discussion of the causative alternation above is idiosyncrasy in the phonological realization of cross-linguistically universal functional heads. So, on root-based views, there are cases in which particular functional heads are genuinely absent, as is the case with particular applicative functional heads in particular languages, and other cases in which they are present, but not phonologically realized. The WYSIWYG view, by contrast, denies the second of these as a source of systematic cross-linguistic variation, and has it that when derivational morphology is systematically not seen, it and any derivation underlying it, is genuinely absent. What I hope to have done in this discussion is to have clarified what some of the predictions are of these contrasting views of the syntax and semantics that underlie derivational morphology. Additional cross-linguistic investigation of these predictions awaits.

CHAPTER 16

ADJECTIVAL AND ADVERBIAL DERIVATION

ANTONIO FÁBREGAS

16.1 SOME RELEVANT CLASSES OF DERIVED ADJECTIVES

In this chapter we will explore some of the analytical issues related to the morphosyntax of adjectival derivation. The first question that we have to consider in order to ground these analytical problems, reviewed in Section 16.2, is what relevant classes of derived adjectives exist and what their main syntactic and semantic properties are. The classification presented in this section is made attending to three factors: (i) the grammatical category of the base; (ii) the kind of notion expressed by the adjective; (iii) the kind of elements the adjective can combine with.

16.1.1 Deverbal Adjectives

Leaving aside stative verbs (but see Rothstein 2004: ch. 7), the main grammatical difference between adjectives and verbs is that prototypically the latter denote events, that is, dynamic processes and changes, while the former are used to express qualities and relations with other entities.¹ This difference can be illustrated if we consider (1),

¹ As is usually the case with very general claims, there are specific empirical phenomena that need to be discussed. The phenomenon of syncategorematicity, by which adjectives like *easy*, *difficult*, *slow*, or *quick* are interpreted as denoting qualities related to some event, is a *prima facie* counter-example. The interpretation of *This kind of book is easy* involves an event (*reading* or *writing*). Note that the event can be expressed as a prepositional complement of the adjective, *easy to read* or *easy to write*; this suggests that the adjective itself expresses here a quality, and the event meaning comes from an implicit complement. Other cases are more problematic: evaluative adjectives, that denote kinds of human

from Spanish. Even though the literature that discusses the properties of participles is too abundant to cover it in a few lines, at least since Wasow (1977) participles are divided in two groups: verbal (1a) and adjectival (1b). See also Levin and Rappaport (1986) for this distinction.

- (1) a. Vi la ciudad [{furiosa-mente/*muy} ataca-da por el ejército
 saw.1SG the city [{furious-ADV/very} attack-PART by the army
 enemiga con bombas de racimo]
 enemy with bombs of cluster]
 ‘I saw the city furiously (*very) attacked by the enemy army with cluster bombs’
- b. La pasta está [{demasiado/*repetida-mente} hervi-da (*por Juan) (*con
 the pasta IS.SL [{too / repeated-ly} cook-PART (by Juan) (with
 una olla)]
 a pot)]
 ‘The pasta is too (*repeatedly) cooked (*by Juan with a pot)’

The participle in (1a) denotes an event, as shown by (i) the possibility of having an agent like *by the enemy army* or an adverbial denoting properties of the agent’s participation, such as *furiously*, (ii) the acceptance of instrumentals such as *with cluster bombs*, (iii) the rejection of degree modification (*very*), (iv) the compatibility with aspectual adverbs like *repeatedly* or *twice*, which count how many times an event takes place. Such participles are frequently embedded under perception verbs, such as *ver* ‘see’, as only events can be perceived (*I saw your nose {turn/*be} red*). The participle in (1b) is adjectival. This is visible in (i) the availability of degree modification (*very*), (ii) the rejection of manner modification, such as *quickly*, (iii) the rejection of agent phrases (*by Juan*), (iv) the rejection of instrumentals (*with a pot*). Adjectival participles are frequent as attributes of copulative verbs, such as *be*.

Verbal participles are generally treated as cases of inflection, because they keep most properties of their verbal base. Beard (1995) has called them transpositions, that is, the result of operations that alter the syntactic distribution of the base without substantially modifying its semantics. The construction shares many relevant properties with the passive, including agent demotion (Lieber 1983), as the participle is predicated from an entity interpreted as a patient of some process. Adjectival participles, on the other hand, are generally treated as cases of derivation: the base’s verbal properties are lost and there is no clear connection with the passive voice. Many adjectival participles are active, that is, the noun they are predicated of is interpreted as the causer or agent of some property, as in Spanish *un libro aburri-do*, lit. ‘a book bor-ed,’ which means ‘a boring book.’ In many cases, the morphological connection with a verb is not matched by the semantics. In *a complicated problem*, we do not interpret that there has been an event

behavior, are frequently interpreted as involving events (*Mary is being cruel* ≈ *Mary is acting cruelly*), but they also allow for modifiers denoting those events (*Mary was cruel to criticize John at the party*) whose syntactic status is not clear (see Stowell 1991, Kertz 2006).

by which someone or something has complicated the problem; we treat that participle as a near-synonym of *difficult*, to denote a quality that does not need to come as the result of a change.

Adjectival participles divide into two main classes: RESULTANT and TARGET state participles (Parsons 1990, Kratzer 2000), here illustrated for English in (2).

- (2) a. The tunnel is {already/*still} completely built (?? by the workers).
- b. The tunnel is {still/*already} completely obstructed (*by the workers).

Example (2a) is a resultant participle: the property expressed by it comes as a result of a building event. The adverbial *already*, which presupposes a past change, can appear, but not the adverbial *still*, which presupposes that the properties can be lost in the future, because once an action has been performed, nothing can change the fact that it was performed, even if the result disappears. Example (2b) is a target participle: the property of being obstructed does not come as the result of an obstruction event. Perhaps the tunnel has never been “non-obstructed,” because the rocks of the mountain have never been removed; hence the availability of *still*, but not *already*, in this reading.

As a consequence of the non-eventivity of adjectives, deverbal adjectives generally have a non-episodic reading where the event expressed by the base is not instantiated, that is, it does not need to take place in actuality for the predication to be true. Three main classes of deverbal adjectives share this property.

- (a) DISPOSITIONAL ADJECTIVES denote the property of being prone to participating in an event. (3) does not denote existing events of forgetting, but the propensity to forget things.
- (3) vergess ‘forget’ > vergess-lich ‘forgetful’ (German)
- (b) POTENTIAL ADJECTIVES are adjectives that express the ability of triggering a particular event. These adjectives are sometimes related to active verbal participles, but differ from them in that the latter entail that an event takes place. Adjectives with *-nte* in Portuguese or some of those derived with *-ing* in English are examples (see (4)). In (4) we denote the property of being able to dissolve, even if it has not been put to work yet.
- (4) solve ‘dissolve, solve’ > solve-nte ‘solvent’ (Portuguese)
- (c) MODAL PASSIVE ADJECTIVES are those that express the possibility or the necessity of undergoing a particular event (Oltra-Massuet 2010). Adjectives built with *-able* in English are typical examples (see (5)). A readable text is a text that can be read easily, even if nobody ever did it. These adjectives are related to middle voice constructions (cf. *Such books typically sell well*), and in languages where this kind of sentence is restricted, these adjectives can be used instead (e.g. Swedish: Klingvall 2008).
- (5) lese ‘read’ > les-bar ‘legible’ (Norwegian)

16.1.2 Denominal Adjectives

The main divide is the one between QUALITATIVE ADJECTIVES and RELATIONAL (or REFERENTIAL) ADJECTIVES. The first (as in (6)) are those adjectives that express properties used to describe entities; they are typically gradable (more about this in Section 16.2.2). The second class (shown in (7)) is formed by adjectives used to classify entities, denoting the domain to which they belong (7a), or to specify other entities with which they establish relations of various kinds, even argumental (7b).

- (6) space > spac-ious, a spacious room (English)
- (7) a. econom-y > econom-ic, an economic problem (English)
- b. Ital-y > Ital-ian, an Italian invasion (English)

Many of the deverbal adjectives, especially those that we called potential, can be assimilated to relational adjectives to the extent that they denote a possible relation with an event. However, it is not standard to assimilate these deverbal adjectives to the class.

Inside qualitative denominal adjectives, several subclasses are generally identified.

- (a) The biggest class, SIMILATIVE ADJECTIVES, is formed by those that denote a resemblance to the notion expressed by the base noun, as in Polish *dziecko* ‘child’ > *dziecinny* ‘childish.’ A typical case is when the adjective denotes the prototypical colour of the base noun, as in English *orange* > *orange*, or when the adjective denotes a resemblance to what is considered to be characteristic of a known person, as in (Dutch) *Dante* ‘Dante’ > *Danteske* ‘Dantesque.’
- (b) QUALITATIVE POSSESSIVE ADJECTIVES are adjectives used to describe an entity by its possession of the notion expressed by the base noun. These adjectives are evaluative. Consider, in Italian, *pancia* ‘belly’ > *panci-uto* ‘with a big belly’. The adjective does not only denote having a belly, but also entails a particular size of it.
- (c) ACTIVITY ADJECTIVES are denominal adjectives used to describe the characteristic behavior of humans. To interpret them, some implicit action is inferred, related to the concept denoted by the base noun: (Swedish) *skoj* ‘joke’ > *skojig* ‘funny,’ that is, when applied to humans, ‘someone that tends to play jokes.’ The implicit action is related to the base noun.
- (d) Related to this last class, some denominal adjectives express the capacity to produce or cause the notion expressed by the base. They will be called here ACTIVE DENOMINAL ADJECTIVES, as in Basque *hidratatze* ‘hydration’ > *hidratatzaile* ‘moisturising,’ as in ‘moisturising cream.’
- (e) CHARACTERISTIC STATE ADJECTIVES express the property of typically being in a state or situation related to the base noun, such as (Catalan) *por* ‘fear’ > *poruc* ‘fearful.’ Here the adjective expresses a ‘passive’ property that the modified entity experiences or suffers, but does not trigger.

Some subclasses of denominal relational adjectives have traditionally been singled out.

- (a) DEMONYMS come from place names and express the relation to a particular territory, typically by birth: (Standard Arabic) *faransa* ‘France’ > *faransi*: ‘French.’
- (b) RELATIONAL POSSESSIVE ADJECTIVES are frequent in Slavic, and they express relations typically marked by the genitive in other languages, without evaluating them, in contrast with qualitative possessive adjectives: Upper Sorbian (Corbett 1987), *bratr* ‘brother’ > *bratrow* ‘brother’s.’

16.1.3 Other Grammatical Categories

Other categories can also produce adjectives, although with a lower productivity. In such cases, the result is frequently a relational adjective. Numerals typically are divided into ORDINALS and PARTITIVES (Finnish: *neljä* ‘four’ > *neljäs* ‘fourth’), which relate an entity to the position it occupies in a series and to a fraction of a whole, respectively. PLACE ADJECTIVES are sometimes derived from locative prepositions and adverbs, as in Spanish *delante* ‘in front’ > *delantero* ‘anterior’ or *tras* ‘behind’ > *trasero* ‘posterior.’

16.1.4 Adjectives from other Adjectives

It is worth noting, finally, that adjectives can be built from other adjectives: *big* > *bigg-ish*, or Spanish *amarillo* ‘yellow’ > *amarill-ento* ‘yellow-ish.’ In such cases, the process is interpreted close to degree morphology: the complex word denotes a lower degree of the property expressed by the base. Something is *biggish* or *yellowish* when it is almost big or almost yellow (cf. Chapter 17 for evaluative affixation).

16.2 ANALYTICAL ISSUES IN ADJECTIVAL DERIVATION

In this section we will examine the main analytical questions in adjectival derivation. In order to provide a detailed picture and to be able to examine fine-grained predictions, we will concentrate on single languages, mainly Spanish and English.

16.2.1 Where Does the Classification Come from?

The first analytical problem related to the previous classes is to determine what the source of this diversity is. Two options suggest themselves: (a) the different classes are due to different adjectivalization processes; (b) the classes emerge because the bases used in each case are different. In (a), the processes themselves would be charged with grammatical and semantic information; in the second, the processes would contribute little more than the grammatical category—TRANSPOSITIONS in Beard's (1995) sense. The answer is not a simple one.

Some affixes behave as predicted by (a). In modern contemporary Spanish, the suffix *-ble* is the clearest case of an affix that comes accompanied by a systematic set of properties, modal passives (8).

- (8) pagable 'payable'; generalizable 'generalizable'; despreciable 'despicable'

The suffix *-ble* is the only productive affix to form adjectives with the meaning 'that {can/must} be X-ed'. One property of *-ble* which is expected if it brings its own semantics is that it imposes its own modal source: some are interpreted as necessary properties (*despreciable* 'that must be despised'), while others denote possible properties (*generalizable* 'that can be generalized'). This contrast does not seem to derive systematically from characteristics of the base.

In contrast, many adjectivalizers behave as expected in (b) (Lieber 2004, Janda 2011). If we attend to the meaning relation that the word establishes with its base, the suffix *-ífico* '*-ific*' produces active denominal adjectives (9a, 'that causes terror'), but also non-active relational adjectives (9b, 'that is related to science') and characteristic states (9c, 'in a state of peace'). The suffix *-oso* is at least used for activity adjectives (10a, *mentiroso*, for instance, means 'that typically tells lies'), active denominals (10b, e.g. *doloroso*, 'that causes pain'), characteristic states (10c, e.g. *gozoso*, 'in a state of joy'), qualitative possessives (10d, e.g. *lacrimoso*, 'that has tears') and similitives (10e, e.g. *cremoso*, 'that looks like cream').

- (9) a. terror-ífico 'terrifying' (< *terror* 'horror')
 b. cient-ífico 'scientific' (< *ciencia* 'science')
 c. pac-ífico 'peaceful' (< *paz* 'peace')
- (10) a. mentir-oso 'lying' (< *mentira* 'lie'); chism-oso 'gossipy' (< *chisme* 'gossip')
 b. dolor-oso 'painful' (< *dolor* 'pain'); grim-oso 'annoying' (< *grima* 'annoyance')
 c. goz-oso 'joyful' (< *gozo* 'joy'); ansi-oso 'anxious' (< *ansia* 'anguish')
 d. lagrim-oso 'weeping' (< *lágrima* 'tear'); call-oso 'calloused' (< *callo* 'callus')
 e. crem-oso 'creamy' (< *crema* 'cream'); leñ-oso 'wood-like' (< *leño* 'log')

In some cases, the meaning of the base helps decide between the different interpretations of the affix. Possessive adjectives take bases denoting physical objects (10d, *montaña*

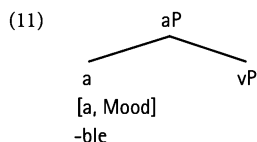
'mountain' > *montañ-oso* 'mountainous'; *pulga* 'flea' > *pulg-oso* 'flea-ridden'). Activity adjectives use bases which denote non-physical entities related to human behavior (10a, *ingenio* 'wit' > *ingeni-oso* 'witty'; *cuidado* 'care' > *cuidad-oso* 'careful'), and characteristic state adjectives are built, as one might expect, over state-denoting nouns (10d, *fervor* 'fervour' > *fervor-oso* 'fervent'; *gloria* 'glory' > *glori-oso* 'glorious'). But the correlation is not perfect. Psychological state nouns can produce active denominals (10b, *fastidio* 'nuisance' > *fastidi-oso* 'annoying') in addition to the characteristic state nouns. Is there any difference between the kind of state expressed by "joy" or "anguish" that makes them different from "peace" and "nuisance"? Not in any obvious way. It would seem, then, that with the suffix *-oso* neither the suffix itself nor the base can directly account for the classes.

It is possible to imagine a proposal where (a) and (b) are combined. This is what some ONOMASIOLOGICAL theories do (Štekauer 2005, 2006). In onomasiological theories, the starting point of the analysis is the correspondence between a morpheme and a particular meaning unit or seme (MORPHEME-TO-SEME ASSIGNMENT PRINCIPLE, Štekauer 2005: 216). Morphemes inside a word are matched to semes which codify the semantics of the concept that the word is intended to express. Cases such as those in (9) and (10) represent situations where the affix can in principle match different semes, given its lexical information, and in such cases the meaning is negotiated depending on its compatibility with the meaning of the base. The affix *-oso* is matched with the meaning "that causes X," "that is in a state of X," or "that has X," among other possibilities, depending on the meaning of the base and its compatibility with each one of these notions. Unlike approaches (a) and (b), the meaning emerges from the interaction of both morphemes, and is not decided by any of the individual affixes, because both have to be matched with the same seme structure.

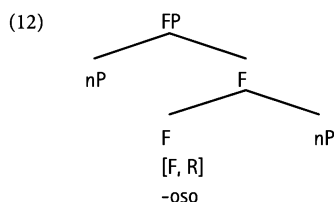
Two options different from (a) and (b) are available in the literature. One option (c) is that the whole word is idiosyncratically associated to a class in a lexical list. Members of pairs like *angustioso* 'stressing' vs. *dichoso* 'joyful' would be listed as units and directly paired with their meaning in a mental dictionary. This approach is particularly compatible with LEXICALISM, specially lexicalist theories that tend to list in the lexicon whole words rather than morphemes (Aronoff 1976, Anderson 1992). Taken to the extreme, solution (c) implies that all words are listed as units, so all classes emerge as the partitioning of the conceptual and grammatical space in the lexicon. This would cause trouble for cases like *-ble*, where there is a close match between morphological marking and class, suggesting that here the affix, not the whole word, is responsible for the properties.

This problem has prompted a fourth solution (d): the classes emerge dynamically as a result of the combination of the units, which is the main strategy in CONSTRUCTIONIST analysis, whose focus is to minimize the information codified in the lexicon and derive as much as possible from the structure in which elements combine. These theories could propose a divide among adjectivizers. Those like *-ble* have a rich feature specification in the syntax (11), as their meaning and grammatical properties are specified. Spanish *-ble* combines only with bases that have a theme vowel, which is a conjugation marker exclusive to verbs. This would be due to a definite selectional restriction imposed by the affix's feature endowment. In the tree in (11), the suffix *-ble* is treated as a head endowed with the features *a* (adjectival) and *Mood*, and it selects a verbal phrase (vP) in order to

produce an adjective derived from a verb; note that *-ble* projects its label, meaning that the whole structure behaves as an adjective.



In contrast, *-oso* would be the exponent of a severely underspecified set of features. This would explain its semantic underspecification: imagine *-oso* is a relational head whose semantics is [R], standing for “relation.” The specific meaning of this relation would depend on the semantic information contained in its syntactic environment and on the pragmatic context. This amounts to saying that *-oso* acts as a preposition or a function marker that relates a set of properties, expressed by its complement, with the subject that holds them. In (12), *-oso* is treated as a head containing the features R and F, for functional, that selects a noun as its complement and defines a relation between that noun and the one merged as its specifier.



Two consequences are expected. The first is that, given its underspecified character, *-oso* must have weak selectional restrictions. Indeed, this suffix creates adjectives from nouns (10), but also from verbal bases, as in *cansa* ‘(to) tire’ > *cansoso* ‘tiresome.’ Secondly, we expect some adjectives in *-oso* to have a vague meaning, as the specific semantic relation is underspecified. This is also borne out. The adjective *arenoso* ‘sandy’ (> *arena* ‘sand’) can be possessive (“with sand”) or similitive (“like sand”). The active and the characteristic state readings are also possible with the same base if the nouns modified by the adjective are different. The adjective *enfadoso* (< *enfado* ‘anger’) can mean “that produces anger,” as in *un trabajo enfadoso* ‘an irritating job,’ but when the subject is an entity that can experience a state, it can mean “that characteristically experiences anger,” as in *una persona enfadosa* ‘a grumpy person.’ However, listing word meanings would still be necessary in cases of adjectives used only in one meaning (cf. *angustioso* ‘stressing’), so this solution is not perfect either.

Before we move to the next section, it must be noted that the core intuition of analysis (d) can be implemented without a hierarchical structure: Janda (2011), inside CONSTRUCTION MORPHOLOGY, develops the idea that some suffixes have no relevant meaning out of context, and their specific meaning emerges through metonymy once used inside a word.

16.2.2 Gradability: Cross-categorial Properties and their Adjectival Instantiation

Prototypically (Croft 1991), adjectives are gradable. We will see, however, that not all of them are gradable in the same way. The big divide is the distinction between qualitative and relational adjectives. Qualitative adjectives (13) generally denote scales, that is ordered sets of values within one dimension, such as “maturity.” They accept INDEFINITE DEGREE adverbs (*very, quite, too . . .*) which select an interval within the scale, located with respect to the standard value that our context defines for a particular domain of comparison, in (13), *for a full professor*.

- (13) John is {very/quite/too/a bit} childish for a full professor.

In contrast, relational adjectives do not denote scales, because they denote relations between entities, and relations exist or do not exist, but do not have degrees. They cannot, thus, accept the degree modifiers in (13) or domains of comparison. What adjectives denoting relations allow are MODIFIERS OF EXTENSION (*completely, partially . . .*). These modifiers indicate whether the relation expressed by the adjective is the only one that the modified noun establishes with an entity, or whether there are other relations that are pertinent in that context. In the first case, modifiers like *completely* are used, while the second meaning is expressed with *partially*. In (14), *completely political* means that the decision entirely falls within the field of politics, and is not related to economy or other factors. To say that the decision is *partially political* means that it does have some relation to politics, but at the same time, is connected to economy, religion or any other domain.

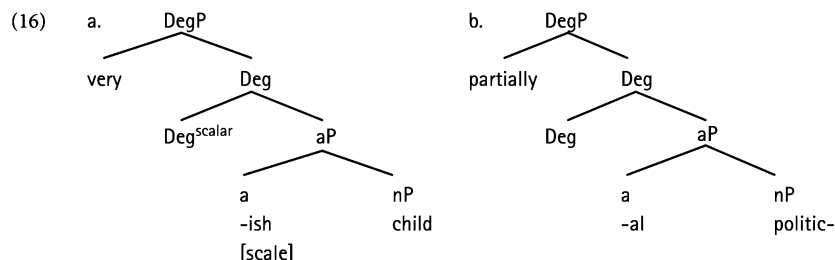
- (14) That decision was {completely/partially} political (*for a constitutional amendment)

Both relational and qualitative adjectives can be part of comparative structures (*This issue is more {economic/problematic} than the previous one*). Gradability is, thus, present in both adjectives, albeit interpreted differently. One reason for the differences can be that relational adjectives are transpositions of nouns, and as such they keep most of the grammatical properties of nouns (Fábregas 2007). Nouns, when used referentially, do not allow indefinite degree adverbs, but they do allow modifiers of extension qualifying the pertinence of the predication (15).

- (15) This object is (partially) a (*very) bed.

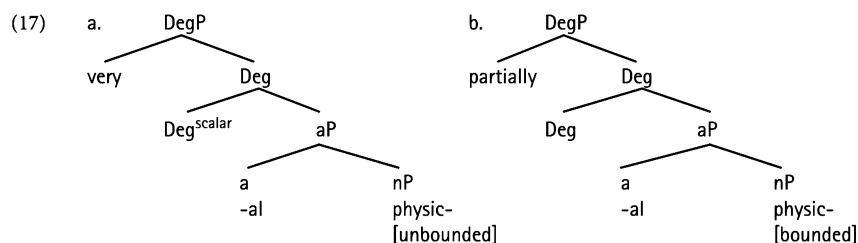
One interesting analytical proposal would be to derive these contrasts from the processes involved. By default, nouns do not express scales, so if the adjective derived from them expresses them, the scale must come from the adjectivizer itself. In (16a), the base

does not denote a scale, but the adjectivizer does; this allows a scalar degree to be projected, licensing an indefinite degree adverb. In (16b), neither the base nor the affix carry scalar meaning, so the degree is not scalar and only extension measurers can be projected (cf. Svenonius (2008) for a possible more complex implementation).



This proposal can be used to explain pairs of relational vs. qualitative adjectives differentiated by affix marking, such as (Spanish) *paterno* ‘paternal’ vs. the similitive *paternal* ‘father-like,’ although it is difficult to systematically associate each one of the readings to single affixes. English does have a contrast between *econom-ic* (relational) and *econom-ic-al* (qualitative), but many adjectives ending in the sequence *-ic-al* are relational (*chemical*, *hydraulical*, *periphrastical*) and some adjectives in *-ic* can be qualitative (*excentric*, *civic*, *eclectic*, *cinematographic*). Another problem are cases where the same affix is used in both members of the pair, as in *physical* (relational: *a partial physical reaction*; qualitative: *A very physical man*).

One option to address these cases is to interpret gradability as a derived concept that comes from a more abstract property. This relates to a central question in the study of grammatical categories: whether categories share some primitive notions which are instantiated as different concepts in different environments (Mourelatos 1978). Scalarity could be derived from BOUNDEDNESS (Jackendoff 1997, Borer 2005b). Boundedness in nouns takes the form of the count ([bounded]) vs. mass (non bounded) distinction, and, in verbs and prepositions, as the telic vs. atelic contrast. In adjectives, it relates to the relational vs. qualitative distinction. This line of research opens the possibility that denominal relational adjectives are built on top of the count readings of nouns, while denominal qualitative adjectives come from their mass readings. That is: *child* in *childish* would be a mass noun (Gallego 2010). The two readings of *physical* would derive from the base noun being taken in a mass noun (17a) or a count noun (17b) reading.



This proposal is, admittedly, difficult to test, as most nouns allow for mass and count readings (Pelletier 1975) and the standard diagnostic tests cannot be applied inside words. Obviously, more needs to be said, and contrasts must be refined.

Let's try to do so. Some scalar adjectives (Hay et al. 1999, Kennedy and McNally 2005) denote CLOSED SCALES, because their scale has a minimal and a maximal value. Adjectival participles like *drunk*, *closed*, or *dressed* have this property. Take *dressed*: there is a maximal degree (= "the whole body is covered") and a minimal degree (= "no part is covered"). Adjectives that lack a maximal value, a minimal value, or both denote open scales, like *dangerous*, *beautiful*, or *fearful*. Modifiers of extension accept only qualitative close scale adjectives, because they denote proportions that must be evaluated within a space defined by two boundaries.

- (18) a. {completely/mostly/half} dressed
 b. {*completely/*mostly/*half} fearful

If we take deverbal adjectives, the prediction of the hypothesis is that atelic verbs produce open scale adjectives, and telic verbs, closed scale adjectives. This is only partially borne out. Stative verbs like *love* (*abundant*, *lovable*, *considerable*, *disgusting*, *resentful*) and activities like *talk* (*talkative*, *navigatable*, *observant*, *attractive*) tend to give open scale adjectives, and accomplishments like *dissolve* (*dissolvable*, *extractable*, *absorbent*, *answerable*), closed scales. In other cases, however, the lack of systematicity suggests that aspectual properties are not always simply translated into scales. Some telic verbs produce open scale adjectives: many *-ing* adjectives (*promise* > *promising*) and dispositionals (*forget* > *forgetful*). With the *-ing* adjectives it is perhaps not entirely implausible to relate the aspectual information with the gerund, which produces atelic readings. In the case of dispositional adjectives, like *forgetful*, it is tempting to pursue the idea that the notion "prone to" acts as an aspectual operator with stative meaning, also atelicizing the base.

In any case, such weakening of the initial proposal implies admitting that the adjectivalization process can modify the aspectual properties of the base, and those modifications are reflected on the behavior of the scale. This is further complicated by two methodological problems: verbs do not always have stable aspectual behavior and the aspectual alternations are not always well-understood (e.g. see Rothmayr (2009) for different ways of producing stative readings of telic verbs). The consequence is clear: attractive as the hypothesis of how to unify gradability, telicity, and countability may be, many more detailed empirical studies are necessary before this is a workable proposal in order to determine, at least, the range of grammatically possible modifications performed by the morphological processes.

16.2.3 Between Inflection and Derivation: Synthetic Comparatives and Superlatives

Morphological degree can be morphologically regular, and in that case it is expressed by specific affixes (example (19)). There are also cases where the the base

is suppletive (20a, b), non-segmentable forms are used (20c) or only one of the forms is segmentable (20d). When the morphological relation between the comparative and superlative is transparent, it can be seen that superlatives are obtained by adding extra material to comparatives (20b, d). This has been interpreted as an indication that superlatives are structurally more complex than comparatives (Bobaljik 2012).

- (19) rich; rich-er; rich-est (English)
 (20) a. good; bett-er; be-st (English)
 b. dobrý; lep-ší; nej-lep-ší (Czech)
 c. bueno; mejor; óptimo (Spanish)
 d. on; hobe; hobe-ren (Basque)

A long-standing question is whether morphological degree is inflectional or derivational. Deciding between the two options is not an easy task. As noted repeatedly (Chapter 2), the traditional distinction between inflection and derivation is difficult to use in practice. Next to prototypical cases of inflection and derivation, natural languages offer many phenomena whose complexity does not allow for a clean adscription to one class. Let us, however, analyze the behavior of degree morphology from this perspective.

Typically, inflectional processes do not change the grammatical category of the word and their semantic contribution tends to be compositional. It is almost non-controversial that an adjective in the comparative or the superlative form still has the properties of an adjective in its syntax and semantics. The adjective *richer* can appear in the same phrases as *rich* and denotes the same set of properties as *rich*, only that in comparison with other entities. However, some cases pose apparent problems to these generalizations. In Spanish, evaluative suffixes can combine with adjectives to express degree (21). In some cases (22), an adjectival base cannot become a noun, but once the evaluative morpheme is there, this becomes possible.

- (21) a. alto
 tall
 'tall'
 b. alt-ito
 tall-dim
 'a bit tall'
- (22) a. delgado > *un delgado
 thin a thin
 'thin' 'a thin one'
 b. delgad-ito > un delgad-ito
 thin-dim a thin-dim
 'a bit thin' 'someone who is (a bit) thin'

However, the form in (22b) can still be used as an adjective. This shows that here degree modification helps to allow the base to undergo adjective-to-noun conversion, but degree does not alter the grammatical category by itself.

It has also been reported that sometimes suppletive comparatives have special semantics: in French (Dietiker 1983), the comparative *pire* 'worse' is used to denote "bad" in abstract situations, as in "being in a bad position," while *plus mauvais* 'more bad' is used in concrete cases, for instance, talking about the behavior of evil people. This is not a counter-example either: the suppletive specializes in a range of the meanings independently allowed by the adjective *mauvais*, but no new semantics is added to the base. There do not seem to be real counter-examples to the previous generalizations, then.

Derivational processes frequently change the Argument Structure of the base, as in the intransitive *wail* vs. the transitive *be-wail*. Degree might seem to alter the items with which the base combines. Comparatives select a SECOND TERM OF COMPARISON introduced with *than* in English (23a); so-called relative superlatives, an EVALUATION SET (23b). The base itself does not combine with these constituents (23c).

- (23) a. Y is a child taller than X.
 b. Y is the tallest child in the class.
 c. Y is a tall child {*than X/in the class}

However, this can be shown to be imposed by the grammatical value added by the degree, and not to involve a change in the lexical properties of the base. The arguments that the adjective takes due to its semantics are never altered by degree morphology. Examples abound: *easy to read* > *easier to read*; *good in maths* > *better in maths*. This shows that the kind of modification performed by degree does not alter the semantics of the property denoted by the adjective, and thus it does not have incidence in the number of participants necessary to fulfill that property. It does, however, license additional modifiers necessary to evaluate a particular degree, in accordance with its grammatical contribution.

Things get more complicated when we consider productivity (Chapter 5). Inflection is prototypically characterized as maximally productive: all words sharing some categorial specification tend to undergo the same inflectional processes. Prototypical derivation is, in contrast, idiosyncratic: words belonging to one category do not undergo the same derivational processes, which might be restricted arbitrarily by properties of individual items. From this perspective, consider the conditions under which English adjectives take morphological or syntactic degree.

The traditional picture is that the availability of the affixes *-er* and *-est* depends on the number of syllables of the base: adjectives of more than two syllables do not take these affixes (24), affixes of one and two syllables take them (25). If this generalization was true, it would not play a role on the morphological productivity of the process: the condition on the number of syllables could be stated as a FILTER on the phonological environment required by *-er* and *-est*, that is, a phonological condition that blocks forms that are otherwise allowed by the grammar.

- (24) difficult > {more difficult/*difficult-er} > {most difficult/*difficult-est}
 (25) happy > {happi-er/*more happy} > {happi-est > *most happy}

However, the data are considerably more complex, and to a certain extent, unstable. The experimental work of Graziano-King (1999) shows that frequency plays an important role in this phenomenon: speakers used the morphological comparative with more frequent items, like *old* and *long*, in 99% of the cases, while less frequent monosyllabic adjectives, like *lax* and *ill*, only took it in 15% of the cases. Bauer et al. (2013) conduct a corpus study and note that the use of the morphological comparative depends on a multiplicity of factors, with frequency being only one of them. We will comment on part of these constraints, those that might be particularly relevant in order to determine if this process is subject to individual lexical preferences, making it closer to derivation.

The first thing that this author notes is that none of the many constraints that play a role in the availability of morphological degree is enough to explain the data. For instance, it is often said that adjectives with a participial form tend to reject the morphological comparative. There is variation inside the group, however. *Stunning* was not documented with morphological degree, but *winning* was; *high-priced* prefers the morphological comparative, but *broad-minded*, which seems to be morphologically identical, prefers the syntactic comparative (Mondorf 2009).

The specific affixes that the base contains also seem to play a role. Adjectives that contain *un-*, *-y*, or *-ly* favor the morphological comparative (26a), while those that contain *-an*, *-ant*, *-ate*, or *-ive* favor the syntactic expression of degree (26b). The examples shown in (26b) were not attested at all with morphological comparative by Bauer et al. (2013).

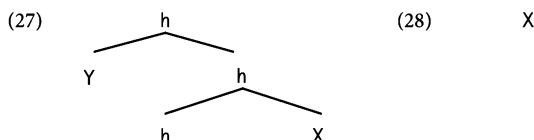
- (26) a. lemony (lemoni-er), unhappy (unhappi-er), friendly (friendli-er)
 b. human, brilliant, private, massive

Semantics also plays a role. Mondorf (2009: 94–5) notes that abstract uses of adjectives, when they are not used in their literal meaning, increase the preference for the syntactic expression of degree. The adjective *fresh*, when used to denote a physical property—*a fresh taste*—only used the syntactic degree in four cases (out of ninety-four attested, slightly more than 4%), while in the metaphorical *a fresh approach* the construction with more was attested in 136 cases (out of 909, almost 15%).

The results are, as with many other morphological operations, not clear. If considered inflectional at all, morphological degree in English is not prototypical at least due to the many idiosyncratic restrictions with individual items. This kind of relatively vague conclusion, frequent in morphology when fine-grained data are considered, has encouraged many researchers to reassess what counts as productive in morphology (Bauer 2001), and some others to dispute that the distinction between inflection and derivation is somewhere codified in the grammar (Marantz 2001).

16.2.4 What is an Adjective? Lessons from Morphology

Let us wrap up the discussion. In the previous short overview, only one thing seems clear: what traditional grammar calls adjective is not defined by a simple list of systematic and homogeneous properties. Not surprisingly, this empirical situation is matched by a great deal of disagreement in theoretical studies about what defines an adjective as a lexical category distinct from verbs and nouns. As an illustration, let us briefly address two different views about the relation between adjectives and nouns. For Hale and Keyser (2002), adjectives are more complex than nouns, that is, categories whose main property is to require a subject of predication. In (27) *X* is the adjective, and *h* is a relational head used to license the adjective's subject, *Y*. If we contrast this structure to the one corresponding to a noun (28), it becomes obvious that for these authors adjectives are structurally more complex than nouns.



Other approaches suggest the contrary view: adjectives are impoverished versions of nouns or verbs. Baker (2003) addresses the issue that not all languages seem to have a grammatically relevant class of adjectives—stative verbs or prepositional phrases typically occupy their place—while verbs and nouns seem to be universally attested. His proposal is that adjectives are defective categories, and that languages with adjectives define them by the absence of properties possessed by nouns or verbs, rather than by positive properties. From this perspective, among lexical categories, adjectives would be the most defective of them and they could be obtained by removing properties from other categories.

The lesson that the behavior of adjectives under morphological operations teaches us is that, perhaps, both views might have to coexist, to the extent that adjectives do not behave in a homogeneous way. As frequently happens in science, it might be the case that the object that we have not been able to fully understand has to be divided in subclasses with distinct properties, and different kinds of structures and processes must be proposed for each subclass (as we did for different kinds of adjectivalizers in Section 16.2.1). Morphology, with the fine-grained sensitivity that individual affixes and processes exhibit, is equipped with precise tools to help diagnose in how many relevant subclasses adjectives should be divided. Probably theoretical linguistics will not be able to understand what underlies the notion of “adjective” without paying attention to its morphological properties and considerable variability.

16.3 ADVERBIAL DERIVATION: RELEVANT CLASSES AND ANALYTICAL ISSUES

Adverbs are a particularly problematic category in linguistic studies. Descriptive grammars notice that they are not a homogeneous class; beyond the traditional requisite that they are uninflected categories, items classified as adverbs inside one language behave differently; compare *not*, *here*, *slowly*, and *perhaps*. Theoretical studies (Chomsky 1965, Jackendoff 1977, Croft 1991, among many others) do not treat adverbs on a par with other lexical categories, and in some cases lack distinctions to endow them with a different feature structure. The general analytical strategy with adverbs has been to analyze them as adjectival or nominal expressions that appear in special structural configurations.

16.3.1 The Relation between Adjectives and Adverbs: Predicative Adverbs

Many adverbs are used to predicate qualities from events, propositions and other concepts expressed by non-nominal syntactic categories, like VP or AP. These PREDICATIVE ADVERBS are frequently built from adjectives (28). Example (28a) predicates from the event denoted by the VP that it is slow; (28b) predicates from the proposition denoted by the whole sentence that it is unfortunate; and (28c) predicates from the speaker of the speech act denoted by the whole utterance that he or she is frank in making that statement. Some adverbs predicate properties from the relation between an event and an argument, such as (28d), where the relation between the agent and the event is described as careful, while some others predicate properties of the subeventive structure of an event, such as (28e), where the adverb expresses that the result obtained in the event was accidental.

- (28) a. Humphrey *slowly* removed the books from the box.
 b. *Unfortunately*, Laura didn't find his keys.
 c. *Frankly*, I do not see your point.
 d. John corrected the exams *carefully*.
 e. Fleming *accidentally* discovered penicillin.

Which entities an adverb can be predicated from is dependent on the properties expressed by the base. Events, but not tables, can be slow, and therefore the adverb in (28a) can be interpreted as a MANNER ADVERB. Whole situations can be unfortunate, so in (28b) the adverb can be a PROPOSITIONALLY-ORIENTED ADVERB; speakers can be frank when uttering something, so *frankly* can be a SPEECH ACT-ORIENTED ADVERB. Agents can be careful in performing an action, but not patients in undergoing it (hence the ungrammaticality of **John fainted carefully*), so *carefully* can be an

AGENT-ORIENTED ADVERB (28d). Similarly, results can be accidental, but not the activities performed volitionally by humans (hence **John accidentally searched for penicillin*), so the adverb in (28e) can be a RESULT ADVERB.

In languages where adjectives morphologically agree with nouns, the relation between the adjective and the adverb generally involves some process that blocks that agreement. As the adverb does not need to agree with a noun, it can combine with grammatical categories different from nouns, such as verbs or whole sentences, that do not carry information about gender or number.

In languages where adjectives agree with nouns, frequently, adverbs derived from them are versions of the adjective where the place of agreement is taken by an invariable marker (29a). If the language has neuter gender, the adverbial form is sometimes identical to the neuter form of the adjective. In Norwegian, when the neuter form of the adjective is marked by *-t*, the same mark is taken by the adverb (29b).

(29)	a.	pulcher	>	pulchr-e (Latin)
		beautiful.MASC		beautiful-ADV
		'beautiful'		'beautifully'
	b.	et klar-t svar ~		å svare klar-t (Norwegian)
		a.NEUT clear-NEUT answer		to answer clear-NEUT
		'a clear answer'		'to answer clearly'

This option is obviously not available in languages where adjectives lack morphological agreement. Languages like English tend to build deadjectival adverbs through extra morphemes. Note in (30), however, that in English the sequence *-ly* is used both to derive adverbs (30a) and to derive adjectives from nouns (30b). One way to interpret this is as a morphophonological accident: there are two homophonous suffixes used for different processes. Another way, coherent with a theory where adverbs do not form a distinct grammatical category, is to consider both as instances of the same affix.

- (30) a. slow > slow-ly
 b. body > bodi-ly

In some European languages these adverbial morphemes are historically related to nouns meaning "manner" or "way." This is the origin of morphemes like *-weise* in German or *-wise* in English.

It is controversial whether adverbs with this marking are now compounds or derived forms (cf. Chapter 3), or even whether the resulting structures can be considered phrasal in some respect. In languages where adjectives inflect, sometimes the adjective in such forms displays something looking a lot like gender agreement, crucially matching the gender that the second morpheme has when used as an independent noun. See (31), from Spanish. In (31a) we find *-mente*, related to the feminine noun *mente* 'mind.' The base appears in the same form when showing feminine gender agreement (31b). It could be argued that in such cases the adjective's agreement is satisfied by the morpheme *-mente*.

- (31) a. clara-mente
clear-ly
b. un-a mente clara
a-FEM mind clear

The same language can have both procedures, the invariable form of the adjective and overt derivation, with differences in the use of each. In Spanish again, the adjective *hondo* 'deep' produces a short adverbial form *hondo* and a derived one, *honda-mente*. The first one specializes in the physical dimension meaning of "deep" and combines with verbs with actions that can define measures in that dimension (32a), while the second takes the metaphorical uses of the adjective, as in "deep thoughts" (32b) and thus tends to combine with mental processes.

- (32) a. cavar {hond-o /*honda-mente}
dig deep-NEUT/ deep-ly
b. pensar {honda-mente/*hondo}
think deep-ly deep-NEUT

An important question in the study of adverbial morphology has been whether the process counts as inflection or derivation. Authors like Sugioka and Lehr (1983) and Bybee (1985) classified deadjectival adverbs as cases of inflection: the adverb is the form in the adjective's paradigm that is chosen in cases where agreement with its subject is not possible. This would make deadjectival adverbs parallel to non-finite verbal forms like infinitives and gerunds. This position has been contested in Scalise (1984) and Zwicky (1995), where the process is treated as derivational.

How can we decide? Note that, as adverbs are not clearly defined as a category, we cannot use the criterion that the process is derivational because it alters the grammatical category of the base. We can, however, rely on other criteria. Adverbial derivation seems to be able to alter the semantics of the base in unpredictable ways. In Spanish, the adjective *seguro* 'certain' is the base of the adverb *segura-mente*, which means 'probably,' not 'certainly.' When we consider productivity, it is clear, again, that not all adjectives in a language produce adverbs. Famously, color adjectives are incapable of this in many languages, like English and Spanish: *red* > **red-ly* (*to paint redly).² With all caveats, this points to a morphological process closer to what has been called derivation.

² There are, however, cross-linguistic differences. Spanish rejects manner adverbs derived from color terms even if the color is interpreted metaphorically. Even though *verde* "green" can mean "ecological," the unavailability of **verde-mente* "green-ly" contrasts with the availability of *ecológica-mente* "ecological-ly." In this kind of metaphorical meaning, where colors are used to denote attitudes, political affiliations, and other notions, Slovak can build manner adverbs from color terms (P. Štekauer, p.c.),

16.3.2 Non-predicative Deadjectival Adverbs: Frame Adverbs

Not all adjectives express qualities. Some relational adjectives can also produce adverbs, but in such cases they are interpreted as FRAME ADVERBS, that is, expressions that restrict the domain to which a particular statement applies (33). What *politically* in (33) states is that, if we restrict our evaluation of the decision to the field of politics, it was wrong; perhaps if we use criteria from other fields, the same decision is right.

(33) Politically, this decision was wrong.

This meaning matches the semantic contribution of relational adjectives: to introduce other entities with which a subject establishes a relation.

Note that the fact that relational adjectives can produce adverbs with the same procedures used by qualitative adjectives implies that these procedures must be sensitive to shared properties between these two classes of adjectives.

16.3.3 Referential or Pronominal Adverbs

Some adverbs have a referential role and can be understood as the adverbial equivalent of pronouns like *he* or *that*. These adverbs are generally underived, but sometimes they are morphologically related to nouns or pronouns. The analysis proposed in Larson (1985) of some denominal adverbs has become almost standard in syntax: referential elements introduced by a (possibly silent) preposition. Along Larson's lines, we could analyse adverbs like those in (34) as nominal constituents—pronouns, nouns, etc.—embedded under prepositional phrases which satisfy their case requisite. The analysis also allows an account of the adverbial uses of nominal expressions like *this Monday* (34c). Again, the intuition is that adverbs are other grammatical categories endowed with some property that satisfies a formal licensing condition that is otherwise active.

- (34) a. I will do it [_{PP} P⁰ to- [-day/-morrow]]
 b. I will do it [_{PP} P⁰ ∅ [now]]
 c. I will do it [_{PP} P⁰ ∅ [this Monday]]
 d. I will put it [_{PP} P⁰ ∅ [here]]

in (i), *pink* meaning “optimistic.” This suggests that there is, indeed, an idiosyncratic restriction in languages like Spanish.

- (i) Vidí svet ružov-o
 see._{3SG} world pink-ADV
 ‘He sees the world optimistically’

Finally, in some languages there are morphological similarities between referential adverbs and pronominal expressions. For instance in English and Russian, some time and place adverbs are perhaps decomposable in a first morpheme shared with pronouns or determiners, and a second morpheme denoting the dimension the pronominal refers to. See Di Sciullo (2005) for an elaboration. This kind of segmentation is controversial. Importantly, phonological changes have to be undergone by the morphemes in order to obtain the right surface forms (cf. the transcriptions).

- (35) a. h-ere (cf. h-e) /hɪər/
 b. th-ere (cf. th-e, th-at, /ðer/
 th-is)

As in the previous cases, this kind of decomposition has to be tested and more detailed empirical work needs to be conducted in a wide variety of languages before it can be accepted or dismissed.

CHAPTER 17

EVALUATIVE DERIVATION

LÍVIA KÖRTVÉLYESSY

17.1 WHAT IS BEHIND THE TERM?

EVALUATION is a mental process by which objects of extra-linguistic reality are assessed from the point of view of quantity (big vs. small) and quality (good, bad, nice, nasty, etc.). Evaluative morphology (EM), in turn, is the means of deriving words that express such concepts. While the concepts of bigness and smallness are measurable, at least in terms of a default value, the concepts of goodness, beauty, or ugliness are, by their very nature, subjective. They concern the highly personal and subtle field of feelings and emotions. Evaluation is an area of morphology in which there is a variety of overlapping and often confusing terminology, so in this chapter we will first take up terminological issues.

Zwicky and Pullum (1987) distinguish between what they call *plain* and *expressive* morphologies. They associate expressive morphology with playful and poetic effects and understand it as a phenomenon “not within the province of the theory of grammar as ordinarily understood. . . the definition of the phenomenon in question lies in domain orthogonal to the grammar” (1987: 9). Dressler and Karpf (1995), however, consider “expressive morphology” to be an inadequate term which they replace with the term *extragrammatical* morphological operations, by which they mean diverse ways of forming new words which deviate from productive rules of word formation. Even though both expressive morphology and extragrammatical morphological operations overlap with evaluative morphology to a certain extent, they focus especially on the qualitative aspect of evaluation. Evaluative morphology, on the other hand, deals mainly with productive morphological rules.

Another term that partially overlaps with evaluative morphology is *appreciative* suffix (Gràcia and Turon: 2000). As the label suggests, appreciative suffixes refer to positive qualitative evaluation. However, Gràcia and Turon’s examples from Catalan fall, no doubt, within the scope of evaluative morphology, for example:

- (1) *gos*
 gos-ic
 ‘small dog’ (Gràcia and Turon 2000: 232)

Russian terminology highlights the subjective nature of evaluative concepts. The term *Suffiksy sub"ektivnoj ocenki* ‘suffixes of subjective evaluation’ (Stankiewicz 1968) refers to suffixes that convey the emotive attitude of the speaker towards the subject of the message. However, as Stankiewicz (1968: 97) points out, “[t]he term ‘subjective’ should not mislead: expressive suffixes are a part of the linguistic code, and their emotive meaning is the same for all speakers; they may, furthermore, signal the emotive meaning independently of the actual emotional state of the speaker.” Inspired by Sapir (1944) Stankiewicz construes the system of expressive derivation as a double axis of polar (“vertical”) and binary (“horizontal”) terms which is implemented by a set of expressive suffixes. The vertical axis represents a graded scale of more vs. less expressive forms which are derived from an emotively neutral base form. On the horizontal axis, expressive forms are opposed to each other as diminutive vs. augmentative and affectionate vs. pejorative. All expressive forms (including diminutives and augmentatives) have an invariant emotive meaning. A similar view is presented by Bauer (1997a: 537) who maintains that the core areas of evaluative morphology are diminutivization and augmentivization although the scope of the field—as he points out—is much broader, including also pejoratives, ameliorative, affectionate, and other connotations.

Finally, the term “expressive” is used by Szymanek (1988). He considers emotions, subjective evaluations, and attitudes “to lie at the foundations of all sorts of ‘expressive’ word-formation” (1988: 106). The categories of diminutiveness and augmentativeness lie, in his view, at “the border area between the cognitively defined core of morphology and its ‘expressive’ periphery” (Szymanek 1988: 106–7). While the categorical content of diminutiveness and augmentativeness dovetails with the bipolar opposition of SMALL–BIG established by the cognitively founded category of DIMENSION, Szymanek points out the differences consisting in the fact that the cognitive concepts SMALL and BIG, respectively, often co-occur with elements of emotional and attitudinal meaning (good, dear, lovely, friendly, etc. vs. bad, hostile, repugnant, etc.). By implication, “the sense-classes identified within the lexicon of ‘expressive’ formations do not merit being regarded as genuine categories of derivation” (Szymanek 1988: 170).

17.2 EVALUATIVE MORPHOLOGY IN LINGUISTIC THEORY

Evaluative morphology has received much attention in the linguistic literature in recent years. In the generative tradition, a prominent issue has been the status of evaluative

morphology as inflection, derivation, or something else entirely. But evaluative morphology has also been studied in terms of its semantic properties, its structural characteristics, and its expression cross-linguistically. In this section we survey some of the prominent literature on the subject.

17.2.1 Inflectional, Derivational or Something Else?

Even very early on Stankiewicz (1968) pointed out the similarity between inflection and expressive derivation. In comparison to non-expressive (lexical) derivation, in both inflection and expressive derivation the meaning of the stem is predictable. Non-expressive derivation “involves invariably a change of the lexical meaning of the stem” (1968: 9). Nevertheless, Stankiewicz follows Trubetzkoy’s (1934) approach and understands expressive derivation as a special case of stem derivation. The status of evaluative morphology as inflection, derivation, or something else entirely has been a matter of contention in more recent scholarship, however, as is briefly outlined in the following sections.

17.2.1.1 Scalise

In a discussion of Italian evaluative suffixes, Scalise (1984) gives a summary of their behavior. Evaluative affixes:

- i. change the semantics of the base (e.g. *lume/lumino*; ‘lamp/little lamp’);
- ii. allow the consecutive application of more than one rule of the same type, and at every application the result is an existent word (cf. *fuoco/fuocherello/fucherellino* ‘fire/little fire/nice little fire’);
- iii. are always external with respect to other derivational suffixes and internal with respect to inflectional morphemes (e.g. *contrabbandierucolo* ‘small time smuggler’ = Word *contrabbando* ‘contraband’) + Derivational suffix (-*iere* ‘agentive’) + Evaluative suffix (-*ucolo* ‘pejorative’) + Inflectional morpheme (-*i* ‘masc, pl.’);
- iv. allow, although to a limited extent, repeated application of the same rule on adjacent cycles (e.g. *carinino* ‘nice+DIM+DIM’);
- v. do not change the syntactic category of the base they are attached to;
- vi. do not change the syntactic features or the subcategorization frame of the base.

He argues that while properties (i) and (ii) match the evaluative suffixes with derivation, (v) and (vi) are typical of inflection. However, properties (iii) and (iv) are neither derivational nor inflectional. Based on this conclusion he suggests a separate block of Evaluative rules within a level ordered morphology (Figure 17.1).

However interesting Scalise’s proposal of the third morphology may seem to be, there is ample evidence that it does not always appear to be valid outside the Italian language for which it was originally constructed. For example, in Supyire (Niger-Congo), evaluative morphology is an inherent part of inflection. Gender 3 in this language is the gender

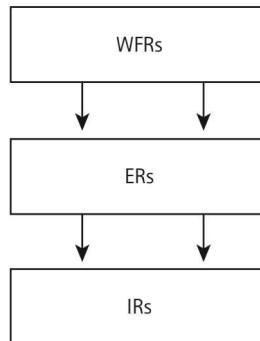


FIGURE 17.1 The place of Evaluative morphology according to Scalise (1984)

of small things. If a root is moved to gender 3 it can denote a smaller object than in another gender (Carlson 1994: 105). Diminutives are results of the change of paradigm:

(2)	Root	Indefinite	Definite	Gloss
	nù-	nòrá	nòróni	'small cow'

In Passamaquoddy (Algonic) the universal derivation vs. inflection order of suffixes is violated, as either order is possible, even in the same form. Diminutive stems may be derived from plural stems and then plurals from diminutives (LeSourd 1995: 126), as example (3) illustrates:

- (3) *hkihka-n-hətə-ss-oltò-kk*
hkihka-n-hətə-ss-ohtò-kk
 all-die-PL-DIM-(A1)-PL-(3)-33AN.ABS
 'all of the poor little ones are dead'

17.2.1.2 *Stump*

Stump (1993) objects to Scalise's model on several grounds. First, he argues that it is not always the case that inflectional rules apply after all word-formation rules. Second, there does not seem to be a justified reason for postulating a separate subcomponent of evaluative morphology. The model does not explain why evaluative rules display precisely the cluster of properties that Scalise claims they do. For example, it is not the case in many languages that evaluative affixes come between derivational and inflectional affixes. In some languages evaluative morphology does change syntactic features or subcategorization.

Stump analyzes evaluative morphology in terms of rules rather than affixes (1993: 12–13). Within his theory of Partial Lexical Rules (1991) Stump classifies rules of

evaluative morphology as belonging to his category-preserving type of derivation and compounding and maintains that as soon as rules of evaluative morphology are covered by category-preserving rules, their properties can be predicted. There is no need assign evaluative rules a peculiar position as a third type of morphology.

17.2.1.3 *Bauer*

Bauer (2004b) adopts Booij’s (1996) distinction between contextual and inherent inflection and assigns evaluative morphology a position between more canonically derivational categories like transposition and categories that verge on being inherent inflection (Figure 17.2).

Although different from Scalise’s position, Bauer’s proposal treats evaluative morphology as distinct and assigns it a unique position. However, he acknowledges the problematic nature of evaluative morphology when he notes that “evaluative morphology is placed awkwardly here, since it is typically class-maintaining, though it can be class-changing” (2004b: 286).

As an interim conclusion, we note that the nature of evaluative morphology depends on a particular language system. Universally speaking, evaluative morphology belongs neither to derivation nor to inflection. Its place is language specific. The evaluative marker changes/modifies the meaning of the base. This can be done by morphological operations which in some languages are closer to the derivation side, and in some others closer to the inflection side of the derivation-inflection continuum.

17.2.2 The Semantics of Evaluation: Jurafsky

While Scalise, Stump, and Bauer discuss the inflectional vs. derivational nature of evaluative morphology, Jurafsky (1996) offers an in-depth view into the semantics of the diminutive. Although he does not deal with the semantics of evaluative morphology in general, his approach is one of the first attempts to cope with the semantic complexities

Contextual	Inherent	Valency-Changing	Evaluative	Transpositional	Lexicon-expanding
Agreement	No agreement				
Lexeme-maintaining		Creating new lexemes			
Class-maintaining			Class-changing		
Grammatical				Lexical	
Paradigmatic					Non-paradigmatic

FIGURE 17.2 An overview of categories by Bauer (2004b)

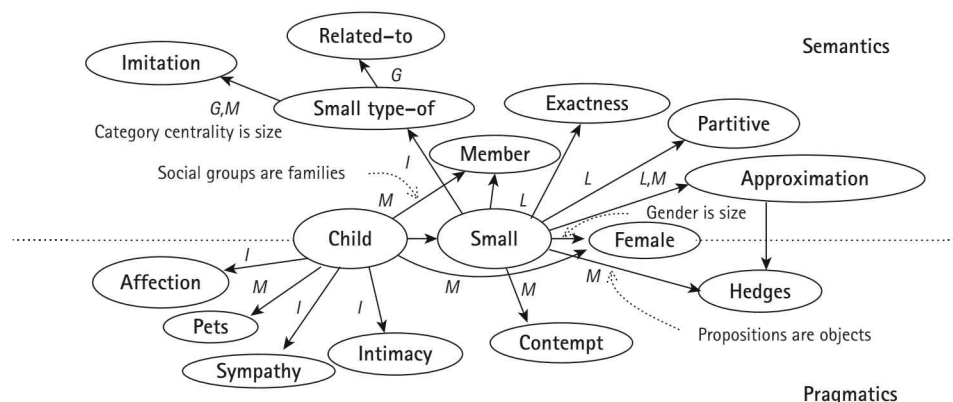


FIGURE 17.3 Synchronic-diachronic model of the semantics of diminutives

of diminutivization. His starting point is Lakoff's notion of RADIAL CATEGORY. Jurafsky models the polysemy of diminutives as in Figure 17.3.

He suggests that the central category of the diminutive is "child." Historically it is the first motivating category of diminutives and it motivates (metaphorically and inferentially) other diminutive senses. Other senses come about through a process of semantic change, for which Jurafsky (1996: 544) distinguishes four mechanisms.

First, semantic change may come about through the creation of metaphors, in this case metaphors for gender or centrality/marginality. With regard to gender, Jurafsky points out two paradoxes. First it is the linking of female gender with both diminutive and augmentative cross-linguistically. Second is the asymmetric use of diminutives and female augmentatives for body parts. The link between female gender and diminutives/augmentatives arises from various sources. In Romance languages, for example, the link results from conflation of the Latin collective suffix with the feminine suffix. In the languages of Southeast Asia, a morpheme originally meaning "mother" has grammaticalized to the augmentative throughout the region (Matisoff 1991). In Indo-European and Afro-Asiatic languages the same morpheme is used for female markers and diminutive markers. At the same time, from the pragmatic point of view, there is a strong tendency for women to use diminutives. The fundamental metaphors for gender INCLUDE ORIGINS ARE MOTHERS; IMPORTANT THINGS ARE MOTHERS (augmentatives); WOMEN ARE CHILDREN; SMALL THINGS ARE WOMEN (diminutives). Metaphors of centrality and marginality are also paradoxical. Diminutive markers can be used to express both intensification (e.g. Mex. Spanish *ahorita* 'immediately') and approximation (Dom. Spanish, *ahorita* 'soon'). Diminutive markers can also mark the centre or the prototype of a social category (Japanese *edo* 'Tokyo,' *edokko* 'Tokyoite') and the social marginal (Fuzhou Chinese *huan-nian* 'foreigner,' where the derogative meaning is introduced by the right-hand diminutivizing constituent). The relevant metaphors include SOCIAL GROUPS ARE FAMILIES (the group member in the source domain corresponds to the child in the target domain); CATEGORY CENTRALITY IS SIZE (it links central or

prototypical members of a category to large size, and peripheral or marginal members of a category to small size); and MARGINAL IS SMALL.

Second, semantic change can come about through the conventionalization of inference. Diminutive markers are naturally associated with inferences such as affection for children and prototype exemplars of small objects. Constructions with such markers can then be lexicalized, as can be seen in the case of classificatory diminutives where an object denoted by a diminutive (for example, *booklet*) is a small object belonging to the same semantic field as the larger object.

A third mechanism of semantic change is generalization or bleaching. An example for this semantic change is the English suffix *-ish* as in *boy-ish*. *Boyish* is less specific than *boy*, and its meaning is more abstract.

A fourth mechanism of semantic change is what Jurafsky calls Lambda-Abstraction Specification which “takes one predicate in a form and replaces it with a second-order predicate, since its domain includes a variable which ranges over predicates” (1996: 555); the outcome of this mechanism is a sense of approximation. The direction in which the diminutive modifies the predicate depends on the direction of the relevant scale. If there is a scale *y* with a point *x* and this point is diminutivized, the resulting meaning is lower than *x* on *y*. For example, in the scale of colors *greenish* is less green than green.

17.2.3 The Onomasiological Approach: Dokulil and Horecký

A different point of departure is taken by the Czech linguist Dokulil (1962). Dokulil represents an onomasiological approach in which the starting point of an analysis of a complex word is meaning rather than form. According to Dokulil there are three basic onomasiological categories: (i) *mutational* (e.g. *write* (ACTION) > *writer* (AGENT)), *transpositional* (*sad* > *sadness*) and *modification* (*duck* > *duckling*). Evaluative morphology falls within the scope of the modification category. In particular, the original meaning is semantically modified, usually enriched. Dokulil (1962: 46) distinguishes four modification sub-categories, illustrated in (4)–(7):

- (4) *Diminutive onomasiological category*—a particular concept is modified by the diminutive marker, e.g. in Hungarian:

házacska
house-DIM
‘small house’

- (5) *Augmentative onomasiological category*—understood as counterpart of the diminutive onomasiological category; a particular concept is modified by an augmentative marker, e.g. Slovak:

domisko
house-AUG
‘big house’

- (6) *Derivation of female names from male names and vice versa*—a concept of animate being is modified by a gender marker, e.g. in Czech:

učitelka

teacher-F

‘teacher’

- (7) *The young of animate beings*—a particular concept is modified by marker with the meaning “not grown-up,” e.g. in English
goose < gosling

Horecký (1964) further develops the idea of the modificational category in reference to the Slovak language. For him, the modificational category serves to express a specific lexical-grammatical modification of a given word. In this way, diminutivizing and augmentivizing morphemes differ from inflectional morphemes and derivational morphemes. Indeed, Dokulil and Horecký seem to have been precursors of Scalise’s third morphology (even if confined to the lexical-semantic facet) which, in fact, only emerged in the generative tradition two decades later.

17.2.4 Recent Approaches

17.2.4.1 *Grandi*

Grandi (2005, 2011a, b) offers both a theoretical background and a comparison of evaluative morphology in various languages. Grandi (2011a) identifies two sides of evaluative morphology: the descriptive, quantitative side represented by diminutives and augmentatives, and the qualitative side that can express a whole range of meanings—reduction/attenuation, intensification, endearment, contempt, authenticity/prototypicality. According to Grandi (2005) an evaluative construction should meet two criteria, one related to semantics, the other to the formal level. From the semantic point of view, the evaluative construction assigns a concept a value that is different from the “standard” value. From the formal point of view, the evaluative construction includes at least an explicit expression of the standard and an evaluative mark. The evaluative mark enriches the base by expressing concepts such as SMALL, BIG, GOOD, BAD. Grandi (2011a) combines the descriptive and the qualitative perspectives and semantic operations assigned to them with a semantic scale that has a positive and a negative pole. From a descriptive perspective, a shift towards the positive end indicates growth, intensification of the actual feature (e.g. physical dimension: *gatto* ‘cat’ > *gattone* ‘big cat’). A shift towards the negative end results in a decrease of the actual feature (*gatto* ‘cat’ > *gattino* ‘small cat’). The combination of the qualitative perspective and a shift towards the positive or negative end indicates the feelings of the speaker. Grandi adopts the suggestion of Wierzbicka (1989: 108) who assumes that “the set of universal semantic primitives must be included in the set of concepts which have been lexicalized in all languages,” and summarizes his ideas in Table 17.1 where SMALL and BIG are semantic primitives representing the

Table 17.1 Descriptive vs. qualitative perspective in evaluation

	Descriptive perspective	Qualitative perspective
Shift towards the "+" end	BIG	GOOD
Shift towards the "-" end	SMALL	BAD

Source: from Grandi (2011a).

descriptive side of evaluation and GOOD and BAD represent the “qualitative” (discourse) side (Grandi 2011a). This approach allows Grandi to identify some prototypes that are supposed to be cross-linguistically constant and recurrent:

- prototypical diminutives indicate a shift towards the negative end on the descriptive axis;
- prototypical augmentatives indicate a shift towards the positive end on the descriptive axis;
- prototypical pejoratives indicate a shift towards the negative end on the qualitative axis;
- prototypical amelioratives indicate a shift towards the positive end on the qualitative axis.

From the point of view of presence of evaluative morphology, Grandi (2011b) divides languages into four types:

- Type A : presence of diminutives; absence of augmentatives;
Type B : presence of both diminutives and augmentatives;
Type C : absence of both diminutives and augmentatives;
Type D : absence of diminutives; presence of augmentatives;

Based on a diachronic comparison of Indo-European languages he observes two different tendencies. First, the original diminutive suffixes have been replaced by new ones—their form has changed, but the semantic function has been preserved. Second, while the Proto-Indo-European lacked augmentative suffixes at all, the majority of present-day Romance and Slavonic languages display a range of various augmentative suffixes. Thus, while diminutivization displays a renewal, augmentivization is a result of innovation.

17.2.4.2 Körtvélyessy

Körtvélyessy (2012) proposes a new approach to the semantics of evaluative morphology in which evaluative morphology is treated as a continuum in which prototypical cases express the meaning of quantity under or above a default value. Körtvélyessy examines evaluative morphology against the background of a “supercategory” of

Quantity that includes not only evaluative morphology but also other categories such as Plurality and Aktionsart, whose concepts of MULTIPLICITY, ITERATIVITY, FREQUENTATIVITY, DISTRIBUTIVENESS, etc., are of quantitative nature (cf. Štekauer et al. 2012). This liberal approach also takes into consideration cases of ATTENUATION (deintensification). The reference point is the standard or default value of the cognitive categories SUBSTANCE, ACTION, QUALITY, and CIRCUMSTANCE. Thus, for example, *blackish* is a diminutive which deviates from the default value *black*. The default value is language specific, influenced by many factors, such as culture, traditions of a speech community, or a speech situation. The key issue of evaluative morphology is the capacity of a language to express morphologically the semantics of “less than/more than the standard quantity,” with the concept of standard quantity being a relative one. The meaning of “other-than-standard” quantity can pertain to any of the aforementioned cognitive categories of SUBSTANCE, ACTION, QUALITY, and CIRCUMSTANCE. By implication, the specific value of standard quantity and any deviations from it may bear on the quantity of both physical and abstract objects, the quantity of actions, processes and events, the quantity of quality and features, and the quantity of particular circumstance. These cognitive categories may be expressed by nouns, verbs, adjectives, adverbs, and also pronouns.

This conception of evaluative morphology can be represented by the following scheme shown in Figure 17.4. As Figure 17.4 illustrates, the process of evaluation starts in extra-linguistic reality. The point of departure is a need in a speech community to evaluate an object of extra-linguistic reality. This need is reflected at the cognitive level. The process of evaluation starts with quantification in terms of the basic cognitive categories (Quantity of Substance, Quantity of Action, Quantity of Quality, and Quantity of Circumstance). If there is a need for qualitative evaluation, based on the iconic semantic shifts SMALL IS CUTE and BIG IS NASTY, the quantitative evaluation can be shifted to qualitative evaluation. For example, in Slovak the diminutive form of ‘cat’ *mačiatko* can refer not only to size, but also to qualities like tenderness, beauty, or cuteness. At the level of the language system, cognitive categories are expressed by semantic categories like diminutive, augmentative, pejorative, ameliorative, plurality, attenuation, intensification, Aktionsart, etc. Concrete realization of these semantic categories comes into existence by means of the markers of evaluative morphology. If needed, the final evaluative construction undergoes phonological changes. The output leaves the level of langue and enters the level of parole where it can obtain various additional shades of emotive coloring, depending on the specific context.

How can this model be projected onto a radial model of EM semantics? In his radial model, Prieto (to appear) aptly replaces Jurafsky’s central category CHILD with the broader category LITTLE. It is proposed here that the central category LITTLE be substituted by an even much broader category of QUANTITY. The reasons are twofold. First, while Jurafsky only proposes a model for diminutives Prieto provides two separate radial models for diminutives and augmentatives. None of them treats diminutives and augmentatives as two central concepts of evaluative morphology in a unified fashion. This comes as a little bit of a surprise because diminutives and augmentatives can be

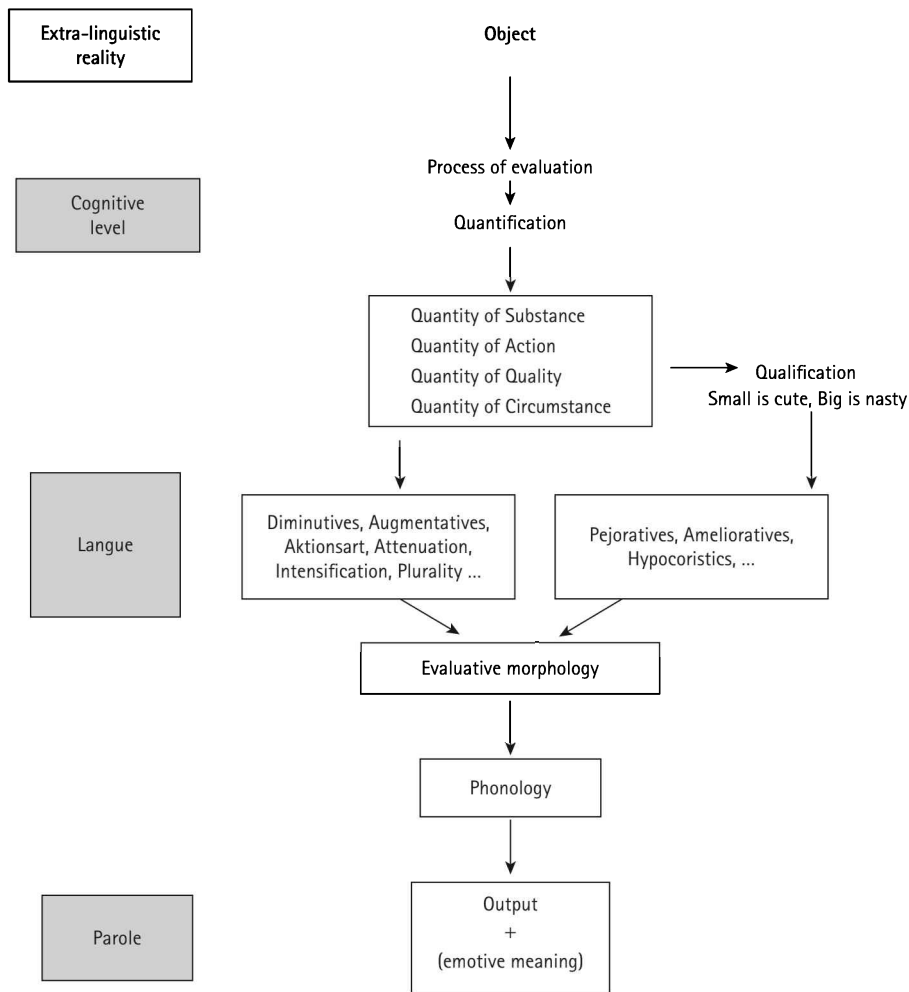


FIGURE 17.4 Model of evaluative word formation

viewed as two poles of a quantitatively defined EM cline, related via a common reference point, in particular, the default value. This default value may be viewed as a prototypical exemplar in regard to quantitative evaluation. Second, any quantitative evaluation with regard to the prototypical default value in the form of various EM meanings (frequentativeness, intensity, duration, distribution, attenuation, exaggeration, approximation, size, social position, etc.) is performed for both diminutives and augmentatives within the scope of the above mentioned cognitive categories of SUBSTANCE, ACTION, QUALITY, and CIRCUMSTANCE. In other words, various EM meanings “radiate” from each of the four cognitive categories. Obviously, individual languages differ in concrete implementation of these cognitively founded options.

All in all, what is proposed here is to take the prototypical, quantitatively defined default value as the central category of a radial model of EM semantics. Diminutives and

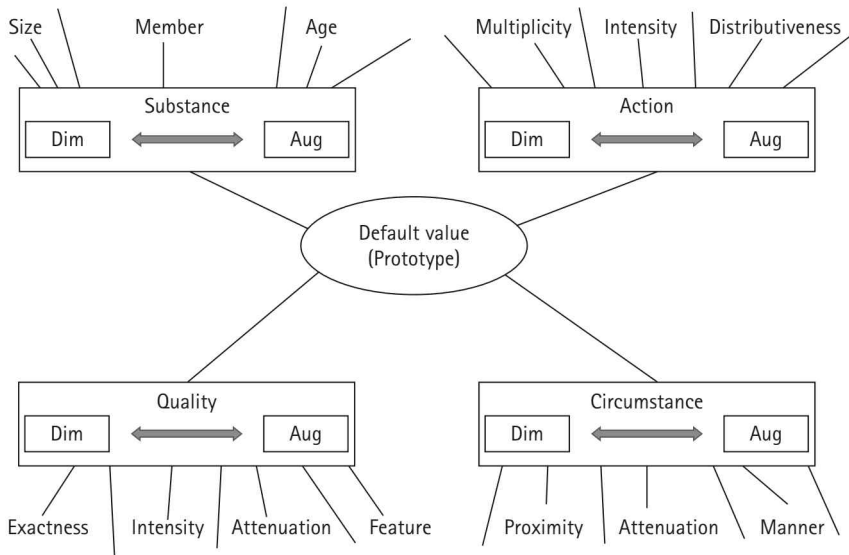


FIGURE 17.5 Radial model of EM semantics

Note: Unlabeled lines indicate other possible meanings.

augmentatives are then viewed as deviations from the prototypical value in any of the cognitive categories (Figure 17.5). It may happen that a particular EM meaning is implemented within two different cognitive categories, such as attenuation which can take the form of a reduced QUALITY (*smallish, reddish*) as well as reduced ACTION (Slovak *skackať* 'to perform very small jumps'—smaller than the prototypical default value).

An interesting sort of support to the unified model of EM semantics comes from cases like the above-mentioned Slovak example *skackať* which indicates not only diminutiveness of ACTION, but also, simultaneously, augmentativeness of ACTION (iterativity—more than the default value of one jump).

Evaluative morphology is not universal, it is language specific. Körtvélyessy (2012) aims to illustrate this through the analysis of a sample of 203 languages. These languages are divided into two groups—the world sample (132 languages) and the Standard Average European languages (71). Furthermore, the languages of the world are divided into six geographical areas (Africa, Eurasia, North America, South America, South-East Asia + Oceania, and Australia and New Guinea). Körtvélyessy introduces a new parameter called *Evaluative Morphology Saturation*. EM saturation is a mean of three values: a word-formation value (V_{WF}), a semantic category value (V_{SC}) and a word class value (V_{WC}). These are numerical representations of the productive use of word-formation processes, semantic categories, and word classes in evaluative morphology in a language. The analysis presented in Körtvélyessy (2012) shows that the most productive word-formation process in the field of evaluative morphology is suffixation, the most frequently expressed semantic category is Quantity of Substance and the most frequently used word class is the class of nouns.

Körtvélyessy also explores possible correlations between evaluative morphology and areal typology, arguing, for example, that the area of Standard Average European languages is characterized by high values of EM saturation in the majority of languages. The next section presents a short survey of evaluative morphology in the languages of the world based on this work.

17.3 EVALUATIVE DERIVATION IN THE LANGUAGES OF THE WORLD

In addition to the important studies discussed above, there are a number of other works that discuss evaluative morphology in general, for example Ettinger (1974), Nieuwenhuis (1985), as well as many publications directly focused on the analysis of diminutives and augmentatives in individual languages: Icelandic (Grönke 1955), North Frisian (Hofmann 1961), Spanish (Gooch 1970, Prieto 2005), Baltic languages (Ambrazas 1993), Hebrew (Bolzky 1994), Passamquoddy (LeSourd 1995), Catalan (Gràcia and Turon 2000), Bulgarian (Derzhanski 2005), English (Schneider 2003), Bickol (Mattes 2006), Sanʿani Arabic (Watson 2006), Slovak (Trnková 1995, Böhmerová 2011), Walman (Brown and Dryer n.d.), German (Ott 2011).

There are also important contrastive works that compare evaluative morphology in several languages, for example: Latin, German, and Romance languages (Ettinger 1980), Polish and Ukrainian (Szymanek and Derkach 2005), English and Slovak (Kačmárová 2010). Fourth, the discussion of diminutives and/or augmentatives is sometimes a by-product of a more central topic, such as Urbanczyk (2006) which discusses reduplicative forms in Central Salish, and Suzuki (1999) who deals with language socialization through morphology in Japanese. Headedness in diminutive Greek word formation is discussed by Melissaropoulou and Ralli (2008), palatalization of bilabials in Xhosa, and Tsonga is studied by Louw (1975).

Pragmatic and morphopragmatic aspects are very often discussed with regard to diminutives. For example, Laalo (2001) analyses diminutives in Finnish child-directed and child speech, Appah and Amfo (2011) discuss the morphopragmatics of the diminutive morpheme (-*ba*/-*wa*) in Akan, Dressler and Barbaresi (1994) devote a whole monograph to morphopragmatics of evaluatives in German, Italian, and other languages, Schneider (2003) includes a pragmalinguistic perspective in his treatment of English diminutives, and Prieto (2005) discusses Spanish evaluative morphology from a pragmatic position, too. Universal tendencies in semantics of diminutives and augmentatives are studied, for example, by Matisoff (1991), phonetic iconicity in evaluative morphology by Gregová (2009) and Panócová (2009). The acquisition of diminutives in thirteen languages is analyzed in a monograph edited by Savickienė and Dressler (2007).

The most recent and most extensive cross-linguist studies of evaluative morphology can be found in Körtvélyessy (2012) and Štekauer et al. (2012). The discussion below is based on Körtvélyessy (2012).

17.3.1 Africa

Suffixation is the dominant word-formation process in diminutivization in the languages of Africa. While the majority of languages make use only of suffixation, there are some languages that use prefixation and compounding to express diminutives. The situation in augmentivization is very similar, showing a strong preference for suffixation and prefixation. The languages of Africa support Jurafsky's theory that the notion of "child" is frequently at the center of diminutive semantics. In Akan (Niger-Congo), for example, the only diminutive morpheme known and documented is the morpheme *-ba*, which is derived from the word *ɔba/ɔbá* 'child' (Appah, pers. com.):

- (8) *a-dɔm-ba*
 a-dɔn-DIM
 'small bell'

Similarly, in Khoekhoe (Khoisan) "child" can be used as a suffix (Chebanne, pers. com.):

- (9) *duu|ua*
 eland.(antelope)-child
 'little antelope'

In accordance with Matisoff's (1991) claims about the semantics of augmentative affixes, in Bafut the prefix *ma-* is used for augmentivization, where *maa* is the regular word for 'mother' (Tamanji, pers. com.):

- (10) *ma-nduu*
 AUG-hammer
 'big hammer'

Examples from languages of Africa are prominent in discussion of the inflectional vs. derivational status of evaluative morphology. Supyire, a Niger-Congo language has been already mentioned (cf. 17.2.1.1). Similar examples can also be found in other Niger-Congo languages. In Aghem, for instance, diminutive meaning is achieved by transfer of nouns from gender class 7/8 to gender class 11/12 (Hyman 1979: 24):

- (11) *fɪ-fú* 'small rat' < *kɪ-fú* 'rat'

Change of paradigm can be also attested in Diola-Fogny. Singular and plural diminutives are formed by means of class 10 and 11 prefixes which are attached to words of different inflectional class, such as -ko 'head' (5/6) and -nil 'child' (1/2) (Aronoff and Fudeman 2005: 62):

- (12) jibekəl/mubekəl 'palm-oil tree'
jikit/mukit 'type of small antelope'

Similarly in Swahili, the noun class prefix ki- (class 7) replaces the plural prefix m- of class 8 (Contini-Morava, pers. com.):

- (13) kitoto 'small child, infant' < mtoto 'child'

17.3.2 Australia and New Guinea

Evaluative morphology is poorly attested in the languages of Australia and New Guinea. Half of the languages in the sample (22 languages) have no evaluative morphology at all and one third of the languages have only diminutives. Languages with only diminutives make use exclusively of suffixation. In languages with both diminutives and augmentatives reduplication is observed. In languages that have only augmentatives, they are expressed through reduplication. In terms of semantics, the diminutive meaning may be secondary to another quantitative meaning. In Bāgandji (Australian), for example, the only diminutive suffix is *-ulu* whose basic meaning is *singular* (Hercus 1982: 81):

- (14) *ɲidja-ulu* *mūrba-ulu*
one child
'one single small child'

17.3.3 North America

The variety of word-formation processes in diminutivization and augmentivization is much bigger in the languages of North America than in Africa or Australia and New Guinea. In North America, we find four formal expressions of evaluative derivation: suffixation, compounding, reduplication, and stem alternation. Cases of stem alternation occur, for example, in Diegueño (Hokan) where voiced laterals alternate with corresponding voiceless lateral spirants. Voiced laterals imply smallness, voiceless laterals bigness or intensity (Langdon 1970: 101):

- (15) *nyily* ‘to be black’ vs. *nyily* ‘to be very black’

With regard to augmentatives, we found no cases of reduplication, but prefixation is used in at least one language, Micmac (Algic) (Hewson 1990: 36):

- (16) *kji-sipu*
 AUG-river
 ‘big river’

17.3.4 South America

The most productive word-formation process in the languages of South America in diminutivization and augmentivization is suffixation. In addition, reduplication, compounding, cliticization, and change in vowel/nasality of a noun are used. Augmentation in Mosestén (Mosetenan) is limited lexically to a group of nouns such as plants or body parts (Sakel 2004: 101):

- (17) *chhi-yijj-si*
 AUG-leg-LNK.F
 ‘big-legged (woman)’

In Tapiete (Tupian), the diminutive suffix *-mi* can modify the lexical meaning of the noun through the formation of a new kinship term (González 2005):

- (18) *shé-si-mi*
 1SG.POSS-mother-DIM
 ‘my maternal aunt’

17.3.5 Southeast Asia + Oceania

The most common word-formation process in this area is also suffixation, but other strategies are employed, such as partial preposing reduplication (19), a combination of prefixation and full reduplication (20), and circumfixation (21):

- (19) Karao *babadiy*
 CV-baliy
 RDP-house
 ‘toy house’ (Brainard, pers. com.)

- (20) Muna *ka-kontu-kontu*
 ka-RDP-stone
 ‘small stone’ (van den Berg 1989: 295)
- (21) Muna *sa-wanu-ha-no*
 CF-get.up-CF-his
 ‘he can barely get up’ (van den Berg 1989: 295)

An analysis of augmentatives in this area reveals two interesting characteristics. First, the Austronesian languages Nêlêmwa and Siar Lak, and the Tai-Kadai language Thai seem to violate what Bakema and Geeraerts (2000: 1046) claim to be an implicational universal, namely that if a language has augmentatives then it will also have diminutives. Second, the most productive way of forming diminutives and augmentatives is not suffixation but reduplication, followed by compounding. Of the ten languages studied in this area, suffixation is found only in Kham (Sino-Tibetan).

The data are, however, in accordance with Matisoff’s (1991) and Jurafsky’s (1993, 1996) claims about the lexical origin of the augmentative and diminutive markers. Coupe (2007: 273) gives an example from the Ao language (Sino-Tibetan). In that language the relational noun *tə-za* ‘child’ (RL-child) has developed into a diminutive suffix and the noun *tə-ji* ‘mother’ (RL-mother) into an augmentative suffix.

17.3.6 Eurasia

Suffixation is the most productive word-formation process in Eurasia with the exception of Malayalam, which relies on compounding. The meaning expressed by Malayalam compounds is the “young one” what is expressed by determinatum *kutti* ‘young’ (Asher and Kumari 1997: 398):

- (22) *pattikkutti*
 dog-young
 ‘puppy’

In Japanese (Bakema and Geeraerts 2000: 1051) the diminutive prefix *ko-* is derived from the noun *kō* ‘child’.

- (23) *ko-same*
 DIM-rain
 ‘small rain, light rain’

Similarly in Ainu, the suffix *-po* originates from the word *po* ‘child, son’ (Refsing 1986: 159):

- (24) *ceppo*
 cep-DIM
 'small fish'

Suffixation is the most productive process in augmentivization, but prefixation is also attested.

17.3.7 Standard Average European

Suffixation occurs in each language with evaluative morphology. In some languages, prefixation, compounding, and/or reduplication are also used. Moreover, Occitan (for diminutives) and Italian (for augmentatives) make use of gender shift, and Maltese relies on the root and pattern technique.

In contrast to the world sample languages, a high number of Standard Average European (SAE) languages make use of more than one word-formation process. In two Slavonic languages, Slovak and Upper Sorbian, four different word-formation processes are used for the coining of diminutives. Derivational processes in the evaluative morphology of SAE languages are richly diversified. Nearly every one of the languages has more than one diminutive suffix. For example, in Polish (Kudła, pers. com.) the suffixes *-k-*, *-ś/ć-*, *-n-*, (among others) are available and can be attached to the same root:

- (25) *piesek*
 pies-ek
 dog-DIM
 'small dog'
- (26) *piesio*
 pies-sio
 dog-DIM
 'small dog'
- (27) *psina*
 pies-ina
 dog-DIM
 'small dog'

Another typical feature of some of the SAE languages is recursiveness, as in the following example from Slovak:

- (28) *malilililinký*
mal<il><il><il>ink-ý
small-DIM-DIM-DIM-DIM-M.SG.NOM
'very veryveryverysmall'

Attaching of more than one diminutive suffix is also possible in Italian (Grandi, pers. com.):

- (29) *bimbettino*
bimb(o)-ett(o)-in-o
child-DIM-DIM-MASC.SG
'small, little child'

All the previous characteristics make the SAE languages unique when compared to the world sample.

17.3.8 Summary—Word-formation Processes in Evaluative Morphology

To sum up, the most productive word-formation process in evaluative morphology is suffixation followed by prefixation, although especially in Germanic languages it is difficult to draw a clearcut borderline between prefixation and compounding. Cases of prefixal-suffixal derivation occur especially in Slavonic languages. Reduplication is the third most productive word-formation process in evaluative morphology with complete, partial preposing, postposing, and infixing being attested. Compounding is used very often in the SAE area as well as elsewhere. Besides these most common word-formation processes, we have also seen examples of stem alternation and change of paradigm. The word-formation processes in evaluative morphology are often accompanied by certain sorts of phonological changes, especially in the process of diminutivization, which is discussed now.

17.4 PHONETIC ICONICITY IN EVALUATIVE MORPHOLOGY

Morphological processes in general are often accompanied by phonological changes and evaluative morphology¹ is no exception. What is especially interesting with regard

¹ For detailed description of evaluative morphology and (mor)phonological changes in diminutives of Indo-European languages cf. Gregová (2009).

to evaluation are the claims that phonological iconicity may be involved. According to Universal #1926 (The Universals Archive) there is a universal tendency for diminutives to contain high front vowels or palatalized consonants and augmentatives to exhibit high back vowels; in other words, tongue position is iconic in the sense that close (or high) equals small, and open (or low) equals large.

It is certainly possible to find examples of this sort of iconicity in individual languages. For example, palatalization can be observed in Polish (Kudła, pers. com.). When two diminutive suffixes are attached, the consonant in the first one is palatalized:

- (30) słów-ecz-ko < słow-o
'little word' word

In Slovak *c* is palatalized to *č*:

- (31) palica < palička
'stick' 'small stick'

Another example comes from Romanian (Štekauer, Valera, Körtvélyessy 2012: 161):

- (32) cărucior < căruță
'pera mbulator' 'cart/waggon'

In Zulu (Bantu) palatalization turns *-d-* in the final syllable into *-j-* in the diminutive (Khumalo and Zeller, pers.com.)

- (33) *isanjwana*
isando-ana
hammer-DIM
'small hammer'

In Algonquian languages the diminutive consonant iconicity is present. It involves the palatalization of **t* to phonetic **[c]* as in the following Cree (Algic) example (Pentland 1975: 241):

- (34) *acihk* or *acihkošiš* 'little caribou' < *atihk* 'caribbou'

However, there seems to be no general cross-linguistic support for this universal. Thus, Ultan (1978) and Nieuwenhuis (1985) propose that phonetic iconicity in morphological diminutives is an areally restricted phenomenon rather than a universal feature of languages. Bauer (1996) analyses fifty-five languages and also concludes that "[t]here does not appear to be any universal principle of sound symbolism operating in markers of the diminutive and augmentative such that palatal articulation correlates with diminutives and not with augmentatives" (1996: 201). Similar results are reported in Štekauer

et al. (2012) for a sample of fifty-five languages. The most comprehensive study in this field was carried out by Gregová et al. (2010). They analyzed twenty-four languages representing three language families (Indo-European, Austronesian, and Niger-Congo). Their research was based on a data sheet consisting of thirty-five lexical items representing core vocabulary and covering four major word-classes: nouns, verbs, adjectives, and adverbs. For the purpose of analysis a three-dimensional model illustrating the place of articulation was applied. As in the earlier studies, their results did not support the existence of the proposed universal. On the contrary, they indicate that front high vowels are typical of augmentatives rather than diminutives and diminutive affixes are phonetically realized by central vowels. At the same time alveolar and velar consonants slightly prevail in both augmentatives and diminutives.

17.5 CONCLUSION

This brief outline of what has been done in the field of evaluative morphology is also a sample of the diversity of topics that fall within the scope of evaluative morphology. Despite the numerous publications (the list above is not complete), a full and comprehensive treatment of the topic has yet to appear.

ACKNOWLEDGMENTS

I wish to express my gratitude to all the informants who kindly supplied the necessary data and answered my questions with endless patience: Pius N. Tamanji, Clement K. I. Appah, Sviatlana Rudaya, Alena Rudenka, Andy Chebanne, Marcela Grygerková, Jan Don, Paul Newman, Nicola Grandi, Sherri Brainard, Agnė Navickaitė-Klišauskienė, Isabelle Bril, Eli Anne Eiesland, Patric Sauzet, Marcin Kudła, Erin Shay, Dumisani Khumalo, Jochen Zeller, Paul Bayethe Damasane.

CHAPTER 18

DERIVATION AND FUNCTION WORDS

GREGORY STUMP

18.1 PRELIMINARIES

IN the canonical case, derivational morphology expresses lexicosemantic relations among lexemes belonging to lexical categories (nouns, verbs, adjectives, and adverbs); accordingly, function words are not ordinarily expected to participate in derivational relations, whether as bases or as derivatives (Naumann and Vogel 2000: 938f, de Groot 2000: 821). Yet, there are clear cases of derivation involving function words. Thus, the question arises: in what ways and to what extent are function words involved in derivation? In order to address this question, one must be clear both about what counts as a function word and about what counts as a derivational relation.

18.1.1 Function Words

Extensional lists of the categories of function words are somewhat variable, but generally include the categories in Table 18.1. This is, of course, an extremely miscellaneous list: some categories of function words inflect (e.g. English personal pronouns and demonstrative determiners), while others do not (English articles and prepositions); some are subcategories of lexical categories (e.g. auxiliary verbs and the proadjective category of French *tel*), some partially intersect lexical categories (some English prepositions also serve as adverbs, others do not—*They ran around*, **They ran into*), and some are disjoint from lexical categories (English coordinating conjunctions and complementizers). Even so, the different categories of function words do exhibit some commonalities.

By definition, FUNCTION WORDS are lexemes which lack independent semantic content, serving instead to mark the grammatical function of an accompanying word,

Table 18.1 Categories of function words

Determiners	
articles	<i>a, the</i>
quantifiers	<i>every, many, some, few, etc.</i>
demonstrative determiners	<i>this, that, these, those</i>
Proforms	
personal pronouns	<i>I/me, we/us, you, etc.</i>
possessive pronouns	<i>my/mine, your/yours, etc.</i>
reflexive pronouns	<i>myself, ourselves, etc.</i> (German <i>mich, uns</i> , etc.; Latin <i>sur</i>)
reciprocal pronouns	<i>each other, one another</i>
indefinite pronouns	<i>somebody, anywhere, no one, etc.</i>
emphatic pronouns	<i>myself, ourselves, etc.</i> (German <i>selbst</i> ; Latin <i>ipse</i>)
demonstrative pronouns	<i>this, that, these, those</i>
relative proforms	<i>who, which, wherever, etc.</i> (Ancient Greek <i>hós</i>)
interrogative proforms	<i>who, what, where, etc.</i> (Ancient Greek <i>tís</i>)
deictic proforms	<i>here, there, now, then, etc.</i>
anaphoric proforms	<i>so in I think so, such in no such thing, etc.</i> (French <i>tel</i>)
Adpositions	
prepositions	<i>into, despite</i>
postpositions	(Turkish <i>gibi</i> 'like,' e.g. <i>bizler gibi</i> 'like us')
Auxiliary verbs	<i>must, be, do in Do you know them?</i>
Particles	
negators	<i>not</i> (French <i>ne, pas</i>)
verb particles	<i>up in wake up, open in rip open, etc.</i>
Conjunctions	
coordinating conjunctions	<i>and, or, but</i>
correlative conjunctions	<i>Either...or, neither...nor</i>
subordinating conjunctions	<i>although, while, since, etc.</i>
conjunctive adverbs	<i>therefore, however, etc.</i>
Complementizers	<i>that, whether, if, etc.</i>
Numerals	<i>one, hundred, thousand, etc.</i>

phrase, clause, or sentence or to modulate its meaning in some systematic way; for example, the article *the* joins with a noun to mark it as definite in reference; the complementizer *whether* joins with a subordinate clause to mark it as interrogative; and the particle *up* has a perfectivizing function when it joins with verbs such as *clean*, *write*,

and *fill*. Function words are thus opposed to CONTENT WORDS, lexemes whose semantic content is largely independent of the particular syntactic context in which they are used. Whereas *Mice!*, *Drive!*, and *Great!* are perfectly meaningful discourse-initial utterances, similar exclamations are unusual or impossible for most function words (**If!* **Into!* **Although!* but *You!*).

An additional feature shared by classes of function words is that they are closed: their membership is small, slow to change and generally invariant across speakers. In this respect, they contrast with such open classes as that of common nouns, whose membership is large and fluid, both across time and across a population of speakers.

Partly as a consequence of their merely adjuvant semantic status, function words frequently exhibit a third, phonological characteristic: they are often unstressed¹ and therefore prone to contraction or cliticization; compare *You will tell them to leave* to [jʌ 'tʰɛɪ əm tʰə 'li:v], whose only unreduced forms are content words.

18.1.2 Derivational Relations

Function words clearly participate in both exocentric and endocentric compounding; for example, the preposition *in* is part of the exocentric compound *shut-in* and of the endocentric compounds *therein* and *inmate* (serving as head in the former case but not in the latter). The concern here, however, is with derivation rather than with word formation generally. Moreover, it is important to distinguish between synchronic derivational relations and diachronic processes of grammaticalization. Consider, for example, the prepositions *during* and *beside*. *During* derives historically from the present participle of **dure* (cf. French *durer* ‘to last’), a verb that no longer exists in English except as part of *endure* and *perdure*. The prepositional status of *during* is the result of a reanalysis of the phrase type *during* NP in sentences such as *During the storm, we remained inside*: originally, *during* NP (also NP *during*) was a nominative absolute construction having NP as its subject (e.g. ‘(with) the storm lasting’), but its gradual disassociation from the obsolescent verb **dure* led to its reinterpretation as a prepositional phrase. Thus, although the function word *during* derives from a verb, it does so through a historical process of grammaticalization (Hopper and Traugott 2003, Narrog and Heine 2011); it does not instantiate any synchronic derivational relation.

Because the noun *side* coexists with the preposition *beside*, *beside* (unlike *during*) can be seen as participating in a synchronic derivational relation. Even so, the derivation of prepositions from nouns is not a productive pattern in English; for example, *beside* ‘at/

¹ Of course, function words are not invariably unstressed; for instance, the pronunciation of polysyllabic function words typically exhibits stress distinctions among their syllables (*mány*, *éither*, *whéther*), and in contrastive or emphatic uses, even monosyllabic function words are stressed (*Look at thém!*). Moreover, a function word may acquire stress by its proximity to a content word, as in Slovak, where all nouns are stressed on the first syllable, from which stress may (under certain prosodic and pragmatic circumstances) pass onto a neighboring preposition. (Thanks to Pavel Štekauer for this observation.)

to the side of' is not paralleled by **becenter* 'at the center of,' nor is *atop* 'on top of' paralleled by **arim* 'on the rim of.' Thus, the derivational relation in which the noun *side* and the function word *beside* participate is merely a relation between existing lexical items—a relation of lexical redundancy. Moreover, this relation is extremely limited in scope, pertaining to only a handful of rather disparate pairs: *(be)side* involves a nominal base, while *(be)low*, *(be)yond*, *(be)hind*, and *(be)fore* involve adjectival bases and *(be)tween* and *(be)neath*, prepositional bases. The unproductiveness of this derivational relation is all the more obvious given that many of its bases are themselves either archaic or otherwise encountered mainly in compounds.²

18.1.3 The Limited Involvement of Function Words in Derivational Relations

The fact that function words serve particular grammatical functions in syntax accounts for their limited role in the sphere of a language's derivational morphology. Languages do not regularly derive innovative grammatical functions; rather, a language's grammatical functions tend to remain constant across time. To the extent that new grammatical functions do emerge, they do so as a gradual effect of language users' reanalyses rather than as the product of synchronic derivation. Thus, there is intuitively no need for the regular derivation of new function words since there is no regular derivation of new grammatical functions. Also, derivation is ordinarily seen as expanding the lexicon, creating one lexeme from another by (i) changing (usually adding to) the phonology of the base lexeme's form (or, in the case of conversion, leaving it unchanged); (ii) changing the syntactic category or subcategory of the base lexeme (or leaving it unchanged); and (iii) changing the lexically specified semantic content of the base lexeme in some way. Because they depend for their interpretation on their syntactic context, function words are not the sorts of things that typically serve as bases for derivation. A language's content words belong to its lexicon; there is a sense in which its function words instead belong to its syntax.

For these reasons, it is unsurprising that an examination of English function words reveals that they mostly fail to participate in derivation, whether as bases or as derivatives. Though they occasionally appear in compounds (*catch-all*, *everyman*), English determiners do not participate in derivation; the same is true of English conjunctions, complementizers, and the negator *not*. Proforms, however, occasionally exhibit cross-formation (Becker 1994) involving a shared demonstrative or interrogative

² The difference between instances of grammaticalization and instances of function-word derivation is not always as clear-cut as the difference between *during* and *beside*. For instance, the preposition-like use of the participle *pending* in sentences like *Pending her decision, they deferred any comment* suggests that it has started down the same path to reanalysis as *during*; but while the finite forms of the verb *pend* have, for many speakers, fallen into obsolescence, the participle still exhibits verbal uses, as in *Her decision is pending*.

component; thus, in archaic English, the adverbial proforms *here/hither/hence*, *there/thither/thence*, and *where/whither/whence* embody an unproductive pattern of cross-formation. Some English prepositions are derived (*beside*, *atop*), and because some prepositions and verb particles also function as adjectives, they are for that reason available to serve as derivational bases; the examples in (1) and (2) illustrate.

(1) *Near*

- a. Prepositional use: *We walked near the lake.*
- b. Adjectival use: *The deadline was nearer than we thought.*
- c. Uses as a derivational base: *near* → *nearly*, *near* → *near* (verb)

(2) *Open*

- a. Use as a verb particle: *We ripped open the box/We ripped the box open.*
- b. Adjectival use: *The thief entered through an open window.*
- c. Use as a derivational base: *open* → *openly*, *open* → *open* (verb)

The auxiliary verbs *do* and *have* are clearly related to the nonauxiliary verbs *do* ‘carry out, execute’ and *have* ‘own, possess (whether alienably or inalienably)’: in their inflection, nonauxiliary and auxiliary *do* share the irregular forms /dʌz/ (* /duz/) and /dɪd/ (* /dʊd/), and nonauxiliary and auxiliary *have* share the irregular forms /hæz/ (* /hævz/) and /həd/ (* /hævd/). In both cases, the relation between nonauxiliary and auxiliary might therefore be seen as a minimally productive relation of derivation by conversion.³ As verbs, the English auxiliary verbs also have derivative gerunds, as in (3).⁴

(3) Gerunds arising from auxiliary verbs

- a. Passive *be*: *Being chosen as chair is an honor.*
- b. Copular *be*: *Being chair requires a lot of work.*
- c. Perfective *have*: *We were surprised at his having taken the job.*

Finally, numerals are exceptional among function words in undergoing various kinds of derivation. For example, the numeral *ten* gives rise to the ordinal adjective *tenth* and its homophonous fractional noun *tenth*; to the multiplicative noun *tenfold*; and to the nominal conversion *ten* ‘a ten-dollar bill’/‘a score of ten’/‘an instance of the numeral 10’).

The tendency for function words to exhibit only limited participation in derivational relations can be seen in all languages with derivational morphology; but some languages

³ Here, “conversion” is intended to refer to any derivational process that is not marked by overt morphology. In this general sense, instances of conversion may or may not involve a change of syntactic category.

⁴ The progressive auxiliary *be* does not have a derivative gerund; this is one manifestation of the so-called Doubl-*ing* constraint (Pullum and Zwicky 1999). The auxiliary *do* has no nonfinite forms because its function is precisely to express “stranded” tense in finite clauses.

Table 18.2 Eight logically possible types of derivational relation involving function words

Base	Derivative	Marking	
content word	function word	base marking	(i, a)
		conversion	(i, b)
function word	content word	base marking	(ii, a)
		conversion	(ii, b)
function word	function word	base marking	(iii, a)
		conversion	(iii, b)
two function words		cross-formation	(x)
a function word and a content word		cross-formation	(y)

show more flexibility than English in allowing certain function words to be derived or to serve as bases of derivation. My purpose here is to examine the nature and extent of this involvement. In particular, I shall examine three questions.

First, what sorts of derivational relations involve function words? Logically, a function word could be involved in any of three sorts of derivational relation: (i) a content word could serve as the base for the derivation of the function word; (ii) the function word could serve as the base for the derivation of a content word; or (iii) the function word could serve as the base for the derivation of another function word. Cross-cutting this three-way distinction among derivational relations is a distinction in morphological marking: a derivational relation may be expressed (a) by **base marking** (the addition of some overt marking⁵ to the base form to produce the derived form) or (b) by simple conversion. Since base and derivative are not clearly distinguishable in instances of cross-formation, I distinguish only two sorts of derivational cross-formation involving function words: (x) cross-formation of two function words and (y) cross-formation of a function word with a content word. The eight logically possible types of involvement of function words in derivational relations are summarized in Table 18.2.

Second, do the derivational relations in which function words participate invariably involve function words (like the relation of derivative prepositions in *be-*), or are there derivational relations in which the participation of function words parallels that of words belonging to lexical categories?

⁵ In most of the examples discussed here, base marking is effected by an affix, but the derivational relations in which function words participate may be marked in other ways as well. Thus, in Georgian, distributive numerals are derived from cardinal numerals by reduplication, e.g. *sami* ‘three,’ *samsami* ‘three apiece’ (Gil 2011, Štekauer et al. 2012: 128); in Southern Vietnamese, tone modulation derives third-person pronouns from family terms (Štekauer et al. 2012: 229); and so on.

Third, are derivational relations invariably unproductive? That is, is the case of the preposition *beside* typical, or do function words sometimes participate in productive derivational patterns?

I consider these three questions in Sections 18.2–18.4. In Section 18.5, I discuss the interactions of derivational morphology with a language's numeral system; numerals deserve separate consideration both because they are atypical as function words and because of the uniqueness of the derivational operations that involve them. I summarize my observations in Section 18.6.

18.2 DERIVATIONAL RELATIONS INVOLVING FUNCTION WORDS

When function words derive synchronically from words belonging to lexical categories, the relation may be expressed by base marking (type (i, a) in Table 18.2), as we have seen in cases such as *beside* and *atop*. Other cases involve simple conversion (type (i, b) in Table 18.2): in the earliest, Vedic form of Sanskrit, the feminine noun *tanū* 'body' may be used as a reflexive pronoun, as in (4); in later Sanskrit, it is supplanted in this function by the masculine noun *ātman* 'soul' (Whitney 1889: §514). In both of these instances, the derivational relation is a relation of lexical redundancy linking a single base to a single derivative. These examples are different from the clearest instances of grammaticalization, since in each of these cases, the base lexeme (with its lexical meaning of 'body' or 'soul') coexists with its derivative (with its anaphoric function as a reflexive proform), and base and derivative are clearly related because they exhibit the same paradigm of forms. But the derivation of function words can instead involve a shift in inflection-class membership. The Sanskrit present participle *bhavat* 'being' may be used as an honorific second-person pronoun (but with third-person verb agreement), as in (5). In this use, its declension differs in one way from that of the ordinary present participle: its nominative singular masculine form is *bhavān*, while that of the ordinary participle is *bhavan* (Whitney 1889: §§452f).⁶ In addition, the derived feminine stem of the honorific pronoun is *bhavatī*, while that of the ordinary participle is *bhavantī*.

⁶ There is some debate about the origin of the honorific pronoun *bhavat*. One view is that it is exactly what it appears to be, an adaptation of the present participle *bhavat* 'being,' which could plausibly have been used in the sense of 'the one being here present' to refer to an honored person. Another possibility is that it is a contraction of the possessive adjective *bhagavat* 'fortunate, blessed' (Whitney 1889: §456); this would explain the long vowel in the nominative singular form *bhavān*, since the inflection of *bhagavat* exhibits this same feature. In any event, its synchronic association is most certainly with *bhavat* (with which it shares all but one of its forms) rather than with *bhagavat* (none of whose forms it shares).

(4)	<i>yajasva</i> worship.2SG. IMPV.MID 'Worship thyself.' (Macdonell 1916: §115)	<i>tanvam.</i> body.F.ACC.SG	(5)	<i>bhavān</i> being.M.NOM.SG 'Your honor goes.'	<i>gacchati.</i> go.3SG.PRES INDIC.ACT

Just as function words may derive from words belonging to lexical categories, the reverse is also true. The function words *over* and *up* serve both as prepositions (*over the rainbow*, *up the ladder*) and as verb particles (*had them over*, *woke them up*): *over* is the base of *overage* 'surplus' (whose affix ordinarily marks denominal or deverbal nouns such as *acreage* and *shrinkage*) and of *overly* 'excessively' (whose affix ordinarily marks deadjectival adverbs such as *excessively*); *up* is the base of the verbal conversion *up* 'raise' (as in *upped the ante*) and of *upward*, which has both adjectival and adverbial uses (*upward mobility*, *looked upward*). These examples embody types (ii, a) and (ii, b) in Table 18.2.

One function word may derive from another. Such derivation may involve affixation (type (iii, a)), as in the case of the deprepositional prepositions *toward* and *unlike* and the gerund of the passive auxiliary in *Being chosen as chair is an honor*. One function word may also arise from another by conversion (type (iii, b)); for example, many verb particles derive from prepositions, in some cases losing any transparent semantic relation to their base (*wipe up*, *freak out*, *do in*, *piss off*). As noted above, proforms may also participate in relations of cross-formation (type (x)). A rather elaborate example of this sort is found in Shughni (an Eastern Iranian language spoken in the Pamir mountains of eastern Tajikistan), which has a number of deictic adverbs related to one another on three dimensions: (i) relative altitude (up, level, down), (ii) distance (here, there, yonder), and (iii) situation (at, near to, motion to); an additional series of adverbs excludes dimension (i), expressing either a locative or an ablative meaning. Formally speaking, the thirty-six adverbs in Table 18.3 embody a completely regular pattern of cross-formation, with each form exhibiting one formative from each of the three sets in (6).

- (6) a. {-*ūd* 'here,' -*ed* 'there,' -*am* 'yonder'}
 b. {— 'at,' -*di* 'near to,' -*ard* 'motion to'}
 c. {*pad-* 'up,' *tar-* 'at the same level,' *ar-* 'down,' *az-* 'at'/'from'}

I know of no case in which a function word participates in a relation of cross-formation with a content word (type (y) in Table 18.2); thus, of the eight logical possibilities distinguished in Table 18.2, only seven seem to be realized.

The fact that referential proforms may serve as derivational bases dispels the notion that words are islands with respect to inbound anaphora (Postal 1969). While it is true that referential personal pronouns do not serve as derivational bases in English (as demonstrated by Postal's example **himite*, cf. *McCarthyite*), that fact must be seen as a local fact about English morphology, since not all languages conform to the same generalization. As Harris (2006) shows, Georgian does not. In Georgian, referential demonstrative

Table 18.3 Deictic adverbs in Shughni

Distance	Situation	Relative altitude			Location/motion away from
		'up' <i>pad-</i>	'at the same level' <i>tar-</i>	'down' <i>ar-</i>	'at'/'from' <i>az-</i>
'here' <i>-ûd</i>	'at' —	<i>pad-ûd</i> 'up here'	<i>tar-ûd</i> 'over here'	<i>ar-ûd</i> 'down here'	<i>az-ûd</i> 'from here'
	'near to' <i>-di</i>	<i>pad-ûd-di</i> 'near to up here'	<i>tar-ûd-di</i> 'near to over here'	<i>ar-ûd-di</i> 'near to down here'	<i>az-ûd-di</i> 'near to here'
	'motion to' <i>-ard</i>	<i>pad-ûd-ard</i> 'up to here'	<i>tar-ûd-ard</i> 'over to here'	<i>ar-ûd-ard</i> 'down to here'	<i>az-ûd-ard</i> 'from here down or into'
'there' <i>-ed</i>	'at' —	<i>pad-ed</i> 'up there'	<i>tar-ed</i> 'over there'	<i>ar-ed</i> 'down there'	<i>az-ed</i> 'from there'
	'near to' <i>-di</i>	<i>pad-ed-di</i> 'near to up there'	<i>tar-ed-di</i> 'near to over there'	<i>ar-ed-di</i> 'near to down there'	<i>az-ed-di</i> 'near to there'
	'motion to' <i>-ard</i>	<i>pad-ed-ard</i> 'up to there'	<i>tar-ed-ard</i> 'over to there'	<i>ar-ed-ard</i> 'down to there'	<i>az-ed-ard</i> 'from there down or into'
'yonder' <i>-am</i>	'at' —	<i>pad-am</i> 'up yonder'	<i>tar-am</i> 'over yonder'	<i>ar-am</i> 'down yonder'	<i>az-am</i> 'from yonder'
	'near to' <i>-di</i>	<i>pad-am-di</i> 'near to up yonder'	<i>tar-am-di</i> 'near to over yonder'	<i>ar-am-di</i> 'near to down yonder'	<i>az-am-di</i> 'near to yonder'
	'motion to' <i>-ard</i>	<i>pad-am-ard</i> 'up to yonder'	<i>tar-am-ard</i> 'over to yonder'	<i>ar-am-ard</i> 'down to yonder'	<i>az-am-ard</i> 'from yonder down or into'

Source: Nekushoeva (2009).

pronouns and deictic adverbs function as bases for a variety of derivatives, some of which are presented in Table 18.4; interrogative, reflexive, indefinite, and personal pronouns can also serve as derivational bases, as in the examples in Table 18.5. The privative circumfix *u-...-o* may even attach to morphological combinations one of whose parts is a referential proform, as in (7). (In the following sections, it will be seen that Sanskrit presents comparable examples of inbound anaphora.)

- (7) *Merab-i k'maqòpil-i-a u-am-c'ign-o-d.*

Merab-NOM satisfied-NOM-is less₁-this-book-less₂-TRANSLATIVE

'Merab is satisfied without these books.' [Literally: '... these — less.']

(Harris 2006: 121)

Table 18.4 Demonstrative pronouns and deictic adverbs as derivational bases in Georgian

Demonstrative pronouns	Proximal	Medial	Remote	Neutral	Gloss
• Possessive forms:	<i>amis</i>	<i>magis</i>	<i>imis</i>	<i>mis</i>	'this' / 'that' / 'yon' / 'his, her'
Derivational affixes					
- <i>tana</i> 'with'	<i>amis-tana</i>	<i>magis-tana</i>	<i>imis-tana</i>	<i>mis-tana</i>	'like this/that'
- <i>iani</i> 'ite'	<i>amis-iani</i>	<i>magis-iani</i>	<i>imis-iani</i>	<i>mis-iani</i>	'her-ite'
- <i>euli</i> 'having once belonged to'	<i>amis-euli</i>	<i>magis-euli</i>	<i>imis-euli</i>	<i>mis-euli</i>	'having once belonged to him'
- <i>eburi</i> 'like, similar to'	<i>amis-eburi</i>	<i>magis-eburi</i>	<i>imis-eburi</i>	<i>mis-eburi</i>	'comparable to this/that'
<i>sa</i> ...- <i>o</i> 'for the purpose of'	<i>sa-amis-o</i>	<i>sa-magis-o</i>	<i>sa-imis-o</i>	<i>sa-mis-o</i>	'for this/that'
<i>u</i> ...- <i>o</i> 'without'	<i>u-amis-o</i>	<i>u-magis-o</i>	<i>u-imis-o</i>	<i>u-mis-o</i>	'without her'
• Modifying forms:	<i>am</i>	<i>mag</i>	<i>im</i>		'this' / 'that' / 'yon'
- <i>it'om</i> 'by reason of'	<i>am-it'om</i>	<i>mag-it'om</i>	<i>im-it'om</i>		'for this/that reason'
Deictic adverbs	<i>ak</i>	<i>mand</i>	<i>ik</i>		'here' / 'there' / 'yonder'
- <i>auri</i> 'originating'	<i>ak-auri</i>	<i>mand-auri</i>	<i>ik-auri</i>		'originating here/there'

Source: Harris (2006: 118).

Table 18.5 Interrogative, reflexive, indefinite, and personal pronouns as derivational bases in Georgian

	<i>vin</i> 'who'	<i>tav-</i> 'self'	<i>aravin</i> 'no one'	<i>me</i> 'I'	<i>šen</i> 'you (sg.)'
Genitive form	<i>vis</i>	<i>tavis</i>	<i>aravis</i>	<i>čem-</i>	
- <i>iani</i> 'ite'	<i>vis-iani</i> 'who-ite'	<i>tavis-iani</i> 'self-ite'	<i>aravis-iani</i> 'no-one-ite'	<i>čem-iani</i> 'me-ite'	<i>šen-iani</i> 'you-ite'
- <i>euli</i> 'having once belonged to...'	<i>vis-euli</i> '... whom'		<i>aravis-euli</i> '... no one'	<i>čem-euli</i> '... me'	<i>šen-euli</i> '... you'
<i>sa</i> ... - <i>o</i> 'for (the purpose of)'		<i>sa-tavis-o</i> 'for oneself'	<i>sa-arvis-o</i> 'for no one'	% <i>sa-čem-o</i> 'for me'	% <i>sa-šen-o</i> 'for you'

Source: Harris (2006).

18.3 THE SCOPE OF DERIVATIONAL RELATIONS INVOLVING FUNCTION WORDS

Some derivational relations serve purely and simply to derive function words. As noted above, some such relations derive function words from functional bases. In Sanskrit, for example, the suffix *-rhi* derives temporal adverbial proforms and *-ti* derives adjectival proforms with a quantificational meaning; in both instances, the derivational bases are pronouns, as Table 18.6 shows. Function-word derivation may also involve lexical bases. An example is the derivation of the English perfect auxiliary from the verb *have* 'own, possess (alienably or inalienably)'. It is clear that speakers of English regard the two verbs as related, since (as noted earlier) they share irregular morphology, but it is also clear that they are grammatically distinct, since they differ in their content and grammatical function (hence in their syntax) and the perfect auxiliary has clitic forms that the main verb *have* lacks; the examples in Table 18.7 illustrate. (The judgments in Table 18.7 are mine.⁷)

Notwithstanding such cases, the participation of function words in a given derivational relation often parallels that of words belonging to lexical categories.⁸ Consider,

Table 18.6 Sanskrit derivative proforms in *-rhi* and *-ti*

Pronominal base		Derivative temporal adverb in <i>-rhi</i>		Derivative quantificational adjective in <i>-ti</i>
<i>amu-</i>	'that'	<i>amurhi</i>	'at that time, then'	—
<i>ta-</i>	'that'	<i>tarhi</i>	'at that time, then'	<i>tati-</i> 'that many'
<i>eta-</i>	'this'	<i>etarhi</i>	'at this time, now'	—
<i>ya-</i>	(relative)	<i>yarhi</i>	'at which time' (relative)	<i>yati-</i> 'as many'
<i>ka-</i>	(interrogative)	<i>karhi</i>	'when?' (interrogative)	<i>kati-</i> 'how many?'

Source: Whitney (1889: §§519, 1103c).

⁷ I'm a native Kansan; speakers of other varieties of English may no doubt disagree with some of these judgments.

⁸ It is important to distinguish instances in which function words participate in the same derivational relation as content words from instances in which derived function words exhibit the same morphological operation as derived content words. In Japanese, for example, reduplication is employed in the derivation of nouns, verbs, adjectives, adverbs, pronouns, adpositions, and numerals (Štekauer et al. 2012: 115), but such derivatives do not all participate in the same derivational relations. For instance, the reduplication of the pronoun *ware* 'I' expresses plural number (*wareware* 'we'), whereas the reduplicated derivative of *gowatsuku* 'be stiff/starchy' is simply a stative adjective (*gowagowa* 'stiff/starchy').

Table 18.7 Differences between the perfect auxiliary and the verb *have* 'own'

	Perfect auxiliary	<i>have</i> 'own'
a. Syntactic differences		
• Inversion	<i>Has John seen it?</i>	* <i>Has John my book?</i>
• Negation	<i>John has not seen it.</i>	* <i>John has not my book.</i> (neutral stress)
• Negative contraction	<i>John hasn't seen it.</i>	* <i>John hasn't my book.</i>
b. Clitic forms		
/v/	<i>I've done it.</i>	<i>I've no idea./?I've an idea.</i>
/z/	<i>He's done it.</i>	* <i>He's no idea./*He's an idea.</i>
/d/	<i>You'd done it.</i>	* <i>You'd no idea./*You'd an idea.</i>
/ə/	<i>I woulda done it.</i>	* <i>I woulda no idea.</i>
/əv/	<i>I would've done it.</i>	* <i>I would've no idea.</i>

for example, the derivational relations marked by the adverb-forming suffixes *-tra*, *-dā*, *-tas*, and *-thā* in Sanskrit: *-tra* (sometimes also *-trā* in Vedic) forms locational adverbs; *-dā* forms temporal adverbs; *-tas*, directional adverbs with an ablative meaning; and *-thā*, manner adverbs. As Table 18.8 shows, each of these affixes joins with pronominal bases and with pronominally declined adjectival bases (many of which have quantificational meanings, and so might be seen as function words); but each also joins with ordinary nominal or adjectival bases. Similarly, the adjective-forming affixes *-vat* and *-īya* in Table 18.9 may take either pronominal or ordinary nominal bases.

A slightly more complicated case of this sort involves the formation of feminine stems from masculine stems in Sanskrit. The pronouns *ta-* 'that', *eta-* 'this', *ya-* 'which (relative)' and *ka-* 'which (interrogative)' and the so-called pronominal adjectives *sva-* 'own', *anya-* 'other', *sarva-* 'all', *viśva-* 'every' and *eka-* 'one' have masculine stems ending in short *a* and feminine stems ending in *ā* (*tā-*, *etā-*, *yā-*, *kā-*, *svā-*, *anyā-*, *sarvā-*, *viśvā-*, and *ekā-*, respectively). This pattern reflects a very general derivational relation in Sanskrit (Whitney 1889: §1149): among ordinary nouns and adjectives, if a masculine stem in *a* has a feminine derivative, this derivative generally has a stem in *ā* (thus, the masculine stems *aśva-* 'horse', *bāla-* 'boy', *bhārya-* 'to be supported (gerundive)', *pāpa-* 'evil' and *priya-* 'dear' have the feminine counterparts *aśvā-* 'mare', *bālā-* 'girl', *bhāryā-* 'wife', *pāpā-* and *priyā-*. But the distinction between masculine *a*-stems and feminine *ā*-stems is crosscut by a second distinction: in the *a*-stem and *ā*-stem declensions, the inflection of pronouns and pronominally declined adjectives differs in certain ways from that of ordinary nouns and adjectives, as the paradigms in Table 18.10 show. Thus, although pronouns participate in the very general derivational relation between masculine *a*-stems and feminine *ā*-stems, they nevertheless maintain their declensional distinctiveness in both categories of stems.

Table 18.8 Some derived adverbs in Sanskrit

	Locational adverb in -tra (also -trā in Vedic)	Time adverb in -dā	Ablative adverb in -tas	Manner adverb in -thā
Pronouns				
Demonstrative ta-	tatra (and V. tatrā) 'there'	tadā 'then'	tatas 'from that'	tathā 'in that manner'
Relative ya-	yatra (and V. yatrā) 'in which place'	yadā 'at which time'	yatas 'from which'	yathā 'in which manner'
Interrogative ka-/ku-	kutra 'where?'	kadā 'when?'	kutas 'from whom?'	kathā 'how, why?'
Pronominally declined adjectives				
sarva- 'every'	sarvatra 'everywhere'	sarvadā 'always'	sarvatas 'from all sides'	sarvathā 'in every way'
anya- 'other'	anyatra 'in another place'	anyadā 'at another time'	anyatas 'from another'	anyathā 'in another manner'
eka- 'one'	ekatra 'in one place'	ekadā 'at the same time'	ekatas 'from one'	cf. ekadhā 'in one way'
Ordinary nouns and adjectives				
samāna- 'same'	samānatra 'on the same spot'			
martya- 'mortal'	V. martyatrā 'among mortals'			
deva- 'god'	V. devatrā 'among the gods'			
puru- 'many'	V. purutrā 'in many places'			
bahu- 'many'	V. bahutrā 'among many'			
nitya- 'continual'		nityadā 'perpetually'		
mukha- 'mouth'			mukhatas 'from the mouth'	
hṛd- 'heart'			hṛttas 'from the heart'	
ṛtu- 'appointed time'				ṛtuthā 'at the appointed time'
ūrdhva- 'rising'				ūrdhvathā 'upwards'

Source: Whitney (1889: §§1098f, 1101, 1103).

Table 18.9 Some derived adjectives in Sanskrit

	Adjective in <i>-vat</i>	Adjective in <i>-īya</i>
Pronouns		
1sg personal pronoun <i>mad-</i>	<i>madvat-</i> , Ved. <i>māvat-</i> 'like me'	<i>madiya-</i> 'my own'
Demonstrative <i>ta-</i>	<i>tāvat-</i> 'so great'	<i>tadiya-</i> 'relating to him/her/it/them'
Relative <i>ya-</i>	<i>yāvat-</i> 'as great'	<i>yadiya-</i> 'relating to whom' (relative)
Interrogative <i>ka-</i> , <i>ki(m)-</i>	<i>kīyat-</i> 'how great?'	<i>kimiya-</i> 'belonging to whom or what?'
Pronominally declined adjectives		
<i>sarva-</i> 'every'	<i>sarvavat-</i> 'containing all'	<i>sarvīya-</i> 'suitable or fit for all'
<i>eka-</i> 'one'	<i>ekavat-</i> 'like one'	<i>ekīya-</i> 'relating to one'
Ordinary nouns		
<i>putra-</i> 'son'	<i>putravat-</i> 'like a son'; 'having a son'	<i>putrīya-</i> 'relating to a son'
<i>Indra-</i>	<i>indrasvat-</i> 'similar to Indra'	cf. <i>indriya-</i> 'belonging to Indra'
<i>grhamedha-</i> 'domestic sacrifice'		<i>grhamedhīya-</i> 'relating to the domestic sacrifice'
<i>svasṛ-</i> 'sister'		<i>svasrīya-</i> 'sister's son'
<i>bhrātṛ-</i> 'brother'		<i>bhrātrīya-</i> 'brother's son'

Source: Whitney (1889: §§451, 517, 1214f, 1233).

18.4 THE PRODUCTIVITY OF DERIVATIONAL RELATIONS INVOLVING FUNCTION WORDS

Because a language's function words ordinarily serve a stable set of grammatical functions and therefore form a closed class, derivational relations that invariably involve function words are in general unproductive: they tend to be relations between existing words rather than relations into which new words may enter. On the other hand, there are productive derivational relations whose base/derivative pairs include pairs whose members belong to lexical categories as well as pairs whose members belong to functional categories. For example, the derivational relation between feminine declensional stems in *-ā* and their masculine counterparts in *-a* is extremely productive in Sanskrit. Even so, its productivity is in general confined to the creation of nominal and adjectival stems that are nonpronominal; it is not typically used for the innovation of new feminine pronouns.

Table 18.10 The distinction between ordinary nouns and pronouns in the *a*-stem and *ā*-stem declensions

		<i>deva-</i> 'god' (m.)	<i>senā-</i> 'army' (f.)	Demonstrative pronoun <i>ta-</i> 'that'	
				(m.)	(f.)
Singular	Nominative	<i>devas</i>	<i>senā</i>	<i>sas</i>	<i>Sā</i>
	Vocative	<i>deva</i>	<i>sene</i>	—	—
	Accusative	<i>devam</i>	<i>senām</i>	<i>tam</i>	<i>Tām</i>
	Instrumental	<i>devena</i>	<i>senayā</i>	<i>tena</i>	<i>Tayā</i>
	Dative	<i>devāya</i>	<i>senāyai</i>	<i>tasmai</i>	<i>Tasyai</i>
	Ablative	<i>devāt</i>	<i>senāyās</i>	<i>tasmāt</i>	<i>Tasyās</i>
	Genitive	<i>devasya</i>	<i>senāyās</i>	<i>tasya</i>	<i>Tasyās</i>
	Locative	<i>deve</i>	<i>senāyām</i>	<i>tasmin</i>	<i>Tasyām</i>
Dual	Nominative	<i>devau</i>	<i>sene</i>	<i>tau</i>	<i>Te</i>
	Vocative	<i>devau</i>	<i>sene</i>	—	—
	Accusative	<i>devau</i>	<i>sene</i>	<i>tau</i>	<i>Te</i>
	Instrumental	<i>devābhyām</i>	<i>senābhyām</i>	<i>tābhyām</i>	<i>Tābhyām</i>
	Dative	<i>devābhyām</i>	<i>senābhyām</i>	<i>tābhyām</i>	<i>Tābhyām</i>
	Ablative	<i>devābhyām</i>	<i>senābhyām</i>	<i>tābhyām</i>	<i>Tābhyām</i>
	Genitive	<i>devayos</i>	<i>senayos</i>	<i>tayos</i>	<i>Tayos</i>
	Locative	<i>devayos</i>	<i>senayos</i>	<i>tayos</i>	<i>Tayos</i>
Plural	Nominative	<i>devās</i>	<i>senās</i>	<i>te</i>	<i>Tās</i>
	Vocative	<i>devās</i>	<i>senās</i>	—	—
	Accusative	<i>devān</i>	<i>senās</i>	<i>tān</i>	<i>Tās</i>
	Instrumental	<i>devais</i>	<i>senābhis</i>	<i>tais</i>	<i>Tābhis</i>
	Dative	<i>devebhyas</i>	<i>senābhyas</i>	<i>tebhyas</i>	<i>Tābhyas</i>
	Ablative	<i>devebhyas</i>	<i>senābhyas</i>	<i>tebhyas</i>	<i>Tābhyas</i>
	Genitive	<i>devānām</i>	<i>senānām</i>	<i>teṣām</i>	<i>Tāsām</i>
	Locative	<i>deveṣu</i>	<i>senāsu</i>	<i>teṣu</i>	<i>Tāsu</i>

Source: Whitney (1889: §§330, 364, 495).

Because derived function words are in general lexically listed, their meaning need not be compositionally computable from their morphology. In English, the derivational suffix *-ward(s)* ordinarily produces denominal adverbs of spatial destination: *backward(s)* 'toward the back,' *eastward* 'toward the east,' *heavenward* 'toward heaven,' *homeward* 'toward home,' *seaward* 'toward the sea'. But with a prepositional base, *-ward(s)* may produce an adverbial derivative having a destinative meaning (*inward* 'toward the inside,'

Table 18.11 Derived pronouns with comparative and superlative morphology in Sanskrit

	Demonstrative pronoun	Relative pronoun	Interrogative pronoun
	<i>ta-</i> 'that'	<i>ya-</i> 'who/which'	<i>ka-</i> 'who/which?'
comparative adjective suffix <i>-tara</i>	<i>tatara-</i> 'that one of two'	<i>yatara-</i> 'who or which of two'	<i>katara-</i> 'which of two'
superlative adjective suffix <i>-tama</i>	<i>tatama-</i> 'that one of many'	<i>yatama-</i> 'who or which of many'	<i>katama-</i> 'which of many'

outward 'toward the outside') or a temporally deictic meaning (*afterward(s)* 'after that', the opposite of *beforehand* 'before that') or may produce a preposition derivative that expresses destination only in combination with a NP complement (*toward(s)* 'in the direction of'). With a verb-particle base, *-ward(s)* produces an adverbial derivative whose meaning hardly differs from that of the particle; compare *We looked up/upward*, *We traveled on/onward*. Even in cases of highly productive derivational morphology, derivatives of function words may have idiosyncratic interpretations. In Sanskrit, for instance, the pronominal stems *ta-* (demonstrative), *ya-* (relative), and *ka-* (interrogative) have derivatives whose morphology is that of comparative and superlative adjectives. The meanings of these derivatives, however, are not like those of ordinary comparatives and superlatives: rather, they serve as pronouns referring to one of two or to one of many individuals, as in Table 18.11.

18.5 DERIVATION AND NUMERALS

Numerals are atypical function words. Indeed, some authors (e.g. Teleman 2005: 1608, Drescher 2009: 96) identify them as content words rather than function words; one reason for this is that in many languages, the class of numerals does not seem closed at all. But a language's simplex numerals (numerals that are not simply combinations of other numerals in an arithmetic relation) are not only always a closed class, but a very small one. English, for example, has fourteen simplex numerals: *one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, hundred, and thousand*. (Learned English also has simplex *googol* for 10^{100} .) Other, complex numerals are ordinarily just morphological or syntactic combinations of these fourteen. Some of these combinations are structurally transparent (*a hundred and one*), while others involve special conjunct alternants; for example, *ten* has the additive alternant *-teen* and the multiplicative alternant *-ty*. Moreover, a language's numerals are diachronically quite stable and tend to exhibit little variability across a population.

Numerals interact with a language's derivational morphology in a distinctive way. First, while compound cardinal numerals are extremely common, most often having either an additive (*thirteen, twenty-five*) or a multiplicative (*thirty, five hundred*) value, cardinal numerals are extremely scarce as synchronically derived forms. In learned English, *-illion* derives names for very large numbers from Latinate bases. Derivatives in *-illion* formerly had different values in American and British English: American English employed (and employs) the so-called short scale, while British English employed the long scale (until 1974, when the United Kingdom officially converted to the short scale). In the long scale, a Latinate base denoting n has an *-illion* derivative denoting 10^{6n} , while in the short scale, a Latinate base denoting n has an *-illion* derivative denoting 10^{3n+3} . For instance, the Latinate base *quadri-* 'four' has the derivative *quadrillion*: in the long scale, *quadrillion* denotes $10^{6 \times 4} = 10^{24}$; in the short scale, *quadrillion* instead denotes $10^{(3 \times 4) + 3} = 10^{15}$. The formative *-illion* is also found in jocular coinages such as *zillion, jillion, bazillion*, and *gazillion*, which masquerade as names for really large numbers, implicating that the speaker regards the number at issue as really large without knowing its precise magnitude.

While cardinal numerals are only rarely derivatives, they serve very widely as bases for the derivation of ordinals. Very commonly, the ordinal forms of the lowest natural numbers stand in a suppletive relation to their cardinal counterparts. Thus, the English cardinals *one* and *two* have the suppletive ordinals *first* and *second*; *third* is less blatantly suppletive. Cardinals above *three* regularly form their ordinals with the derivational suffix *-th* (with laxing of /a/ to /ɪ/ in *fifth* but not *ninth*); nevertheless, compound cardinals ending in *one, two*, or *three* have compound ordinals ending in *first, second*, and *third*, e.g. *twenty-first*.

Languages vary widely in the manner in which ordinal morphology is employed in the derivation of compound ordinals. In Kanuri (Nilo-Saharan), the ordinal derivative of a compound cardinal of form X arises through the circumfixation of *kán-* ... *-mi*, as in Table 18.12a. In English, the ordinal derivative of a compound cardinal of the form [Cardinal₁ Cardinal₂] (where Cardinal₂ is a simplex numeral) exhibits ordinal morphology on Cardinal₂ only, as in Table 18.12b. In Finnish, a compound cardinal of the form [Cardinal₁ Cardinal₂] may have an ordinal derivative in which both Cardinal₁ and Cardinal₂ exhibit ordinal morphology, as in Table 18.12c. In Modern Greek, the morphology of the ordinal derivative of a compound cardinal of the form [Cardinal₁ Cardinal₂] depends on whether the relation between Cardinal₁ and Cardinal₂ is multiplicative or additive: in the former case, only Cardinal₂ exhibits ordinal morphology; in the latter case, both Cardinal₁ and Cardinal₂ do, as Table 18.12d shows. In Breton, the ordinal derivative of a compound cardinal of the form [Cardinal₁ Cardinal₂] exhibits ordinal morphology on Cardinal₂, but the ordinal derivative of a compound cardinal of the form [Cardinal₁ and Cardinal₂] instead exhibits ordinal morphology on Cardinal₁; the examples in Table 18.12e illustrate. Breton additionally exhibits a common phenomenon in ordinal marking: a cardinal numeral with a suppletive ordinal in isolation may exhibit a regular ordinal in combinations. Thus, in isolation, *unan* '1' has the suppletive ordinal *kentañ* 'first', but the ordinal form of the compound cardinal *unan warn-ugent*

Table 18.12 Ordinal derivatives of compound cardinal numerals in five languages

	Language	Compound cardinal		Ordinal
a.	Kanuri	<i>findin</i> 20	<i>(lúkko)</i> <i>tilón</i> '21' 1	<i>kân-findin</i> <i>(lúkko)</i> <i>tilón-mi</i> '21st' ORD-20 and 1-ORD
b.	English	twenty-one		twenty-first
c.	Finnish	<i>kolme-tuhatta</i> 3-1000.PART.SG	<i>sata-kolme-kymmentä-neljä</i> '3134' 100-3-10.PART.SG-4	<i>kolmas-tuhannes</i> <i>sadas-kolmas-kymmenes-neljäs</i> '3134th' 3.ORD-1000.ORD 100.ORD-3.ORD-10.ORD-4.ORD
d.	Modern Greek	<i>xília</i> <i>tetrakósia</i> 1000 4.100	<i>peninda</i> <i>téssera</i> '1,454' 5.10 4	<i>xíliostós</i> <i>tetrakosiostós</i> <i>pendikostós tétartos</i> '1,454th' 1000.ORD 4.100.ORD 5.10.ORD 4.ORD
e.	Breton	<i>pem-zeg</i> '15' 5-10	<i>pemp</i> <i>ha</i> <i>tri-ugent</i> '365' <i>tri</i> <i>c'hant</i> <i>pemp</i> <i>ha</i> <i>tri-ugent</i> '365th' 3 100 5 and 3-20	<i>pem-zeg-ved</i> '15th' 5-10-ORD <i>tri</i> <i>c'hant</i> <i>pemp-ved</i> <i>ha</i> <i>tri-ugent</i> '365th' 3 100 5-ORD and 3-20

Source: Stump (2010 and sources cited therein).

‘21’ (literally ‘1 on-20’) is *unanved warn-ugent* (where *unan* takes the regular ordinal suffix *-ved*) rather than **kentañ warn-ugent*. Compare also French *un* ‘1’ with suppletive ordinal *premier* ‘1st’; but *vingt-et-un* ‘21’ with regular ordinal *vingt-et-unième* ‘21st’; not **vingt-premier*. (See Stump 2010 for additional discussion.)

Ordinal morphology is often versatile in its function. In English, ordinals also serve as fractionals (e.g. *fifth* ‘1/5’), except in the case of *second* (whose fractional counterpart is *half*). In Sanskrit, the ordinal suffix *-tha* (*caturtha-* ‘fourth,’ *ṣaṣṭha-* ‘sixth’) also joins with the quantificational proforms *tati-* ‘that many,’ *yati-* ‘as many,’ and *kati-* ‘how many?’ (see Table 18.5) to produce quantificational ordinal proforms: *tatitha-* ‘the that-manieth,’ *yatitha-* ‘the as-manieth,’ *katitha-* ‘the how-manieth?’ (These forms were apparently reanalyzed as combinations of the pronominal stems *ta-*, *ya-*, and *ka-* with a suffix *-titha*, which in turn joined with *tāvat-* ‘so great’ and *yāvat-* ‘as great’ (Table 18.8) to produce *tāvatitha-* ‘the that-manieth’ and *yāvatitha-* ‘the as-manieth’ and with the quantificational adjective *bahu-* ‘much, many’ to produce *bahutitha-* ‘manifold’; Whitney 1996: §1242d.)

In some instances, such versatility suggests that ordinal morphology is a specialization of morphology that once had a different function. In many Indo-European languages, there is a well known link between the morphology of superlative adjectives and that of ordinal numerals (Bopp 1833: 457ff). In Sanskrit, for example, *-tama* is one of the markings used to form ordinal stems for numbers beginning with 20: *viṃśati-* ‘20’ → *viṃśatitama-* ‘20th.’ This same affix is one of the markings used to form superlative adjective stems: *priya-* ‘dear’ → *priyatama-* ‘dearest.’

Cardinal numerals also commonly serve as bases for the derivation of numeral adverbs (e.g. Latin *septiēns* ‘seven times,’ *octiēns* ‘eight times’) and of distributive adjectives (*septēni* ‘seven each,’ *octōni* ‘eight apiece’), often with suppletion for lower numeral bases (Latin *semel* ‘once,’ *bis* ‘twice’; *singuli* ‘one apiece,’ *bīni* ‘two each’). In English, both native and Latinate numeral bases serve in the derivation of multiplicatives (*fourfold*, *quadruple*); these may have adjectival, adverbial, and verbal uses (*a threefold increase*; *a triple helping of potatoes*; *The donations increased threefold*; *They tripled the amount*). Both native and Latinate bases are also used for the derivation of numeral collective nouns, often with idiosyncratic narrowing of meaning; for example, *foursome* and *quartet* both refer to groups of four people, but the former is conventionally used to refer to two competing pairs of golfers while the latter is conventionally used to refer to four musicians who perform together.

Cardinal numerals, alone or in combinations, are frequently converted to lexemes belonging to one or another lexical category—mainly (though not exclusively) to common nouns referring to things whose value is measured in numbers or whose use or appearance in some way involves a particular number; Table 18.13 gives some representative examples from American English. Comparable examples appear in many languages. In discussing naming practises in ancient Rome in his *Classical and archaeological dictionary of the manners, customs, laws, institutions, arts, etc. of the celebrated nations of antiquity and of the Middle Ages*, Nuttall (1840: 313) notes that

Table 18.13 Content words derived from numerals, mostly by conversion (American English)

Derivative	Meaning	Usage example
<i>one</i>	noun: \$1 bill	<i>three ones</i>
<i>five</i>	noun: a) \$5 bill b) a handshake or congratulatory hand slap	<i>three fives</i> <i>Gimme five!</i>
<i>deep-six</i>	verb: reject	<i>Deep-six him!</i>
<i>seven-eleven</i>	a) proper noun: name of a chain of convenience stores b) common noun: a convenience store belonging to that chain	<i>She owns stock in Seven-Eleven.</i> <i>three Seven-Elevens</i>
<i>nine-eleven</i>	proper noun: a) September 11, 2001 b) terrorist attack in New York City on that date	<i>Was 9/11 preventable?</i>
<i>ten</i>	noun: a) \$10 bill b) person or thing rated at 10 on a scale of 1 to 10	<i>three tens</i> <i>S/he's a ten.</i>
<i>teen</i>	noun: teenager (derived by clipping rather than conversion)	<i>three teens</i>
<i>twenty</i>	noun: \$20 bill	<i>three twenties</i>
<i>twenty-two</i>	noun: .22 caliber handgun	<i>three twenty-twos</i>
<i>twenty-four seven</i>	adverb: all the time (i.e. 24 hours a day, 7 days a week)	<i>They were on duty 24/7.</i>
<i>thirty-eight</i>	noun: .38 caliber handgun	<i>three thirty-eights</i>
<i>forty-five</i>	noun: a) vinyl recording played at a speed of 45 rpm b) 45 caliber handgun	<i>three forty-fives (ambiguous)</i>
<i>fifty</i>	noun: \$50 bill	<i>three fifties</i>
<i>seventy-eight</i>	noun: vinyl recording played at a speed of 78 rpm	<i>three seventy-eights</i>
<i>eighty-six</i>	verb: cancel, reject	<i>Eighty-six that order!</i>
<i>eighty-eight(s)</i>	noun: a) butterfly in the genus <i>Diaethria</i> (with an '88'-like pattern on its wings) b) (definite plural only) piano	<i>three eighty-eights</i> <i>play the eighty-eights</i>
<i>hundred</i>	noun: \$100 bill	<i>three hundreds</i>
<i>one eighty</i>	noun: a full reversal	<i>They did a one eighty.</i>

[a]n only daughter took the family name alone, which was frequently softened into a diminutive, as Tulliola, for Tullia. If there were two, three, or more daughters, they were distinguished by a word expressive of seniority, &c., or by numerals. Where there were two, for instance, they were distinguished by the addition of senior and junior, or major and minor; where there were three or more, they were called Prima, Secunda, Tertia, Quarta, Quinta, &c., or these ordinals were made diminutives, as Secundilla, Quartilla, Quintilla, &c.⁹

18.6 SUMMARY

In general, categories of function words participate only marginally in a language's synchronic derivational relations. In a given language, certain categories of function words may altogether fail to serve either as bases of derivation or as derivatives. If a category of function words does participate in derivation, the relation involved is very often unproductive, relating only a small number of existing lexical items. Nevertheless, there are clear instances in which words belonging to a functional category parallel words belonging to a lexical category in their participation in a particular derivational relation; such instances may even involve productive derivational relations. In general, proforms, adpositions, auxiliaries, verb particles, and numerals appear to participate in derivational relations more frequently than determiners, conjunctions, and complementizers.

A derived function word may arise from either a content word (e.g. *side* → *beside*) or another function word (e.g. Sanskrit *ka-* (m.) 'who? what?' → *kā-* (f.) 'who? what?'), and a function word may serve as base for the derivation of a content word (e.g. *over* → *over-age*) or another function word. The derivational relation in which a function word participates may be expressed by affixation (e.g. Sanskrit *ka-* 'who? what?' → *katara-* 'which of two?'), by conversion (e.g. *do* 'carry out, execute' → auxiliary *do*) or by cross-formation (e.g. *whence/whither* ↔ *hence/hither* ↔ *thence/thither*).

The involvement of particular categories of function words in derivational relations varies cross-linguistically. For example, referential proforms do not generally serve as derivational bases in English, but they do in others (as in Georgian and Sanskrit); this fact demonstrates that words do not act universally as islands for purposes of inbound anaphora.

⁹ On the formation of diminutives from numerals, see further Bauer (1997a: 540 ff).

CHAPTER 19

POLYSEMY IN DERIVATION

FRANZ RAINER

19.1 INTRODUCTION

THOUGH multiplicity of meaning is no less common in patterns of word formation than it is in simplex words, it has rarely been treated systematically on its own terms (Haspelmath and Sims 2010, for example, has no item *polysemy* in the index; cf., though, Lehrer 2003, Lieber 2004, Rainer 2005a, 2010). The reason for this relative neglect probably has to be sought in the—generally tacit—assumption that polysemy in word formation follows the principles that underlie polysemy in general and that therefore there is not much to say about word formation. One of the few studies specifically dedicated to the problem at hand indeed reaches the conclusion that “[a]lthough the concepts represented by English affixes are limited, they show polysemy that is similar to that found in lexemes” (Lehrer 2003: 229). I will try to show that nevertheless the issue of polysemy in word formation is worth studying on its own.

19.2 BASIC NOTIONS

The concepts “homonymy” and “polysemy” reach back to Ancient Greece, where words with multiple senses were already discussed as a problem of rhetoric and philosophy (cf. Kratschmer 1996, Engels 2003). In the wake of Aristotle, ambiguity was generally considered as a vice until well into the modern times, though it was not denied that it could also yield interesting stylistic effects under special circumstances. In contrast to this rhetorical/philosophical tradition, modern linguists tend to emphasize the positive side of words having multiple senses, a feature of natural languages that is held to underlie their incomparable flexibility (cf. Nerlich and Clarke 2003) and to notably facilitate memorization, while misunderstandings—the reason for the Aristotelian anathema against the multiplicity of senses—have been found to be relatively rare in

real communication thanks to the disambiguating role of context and background knowledge.

As far as the terms themselves are concerned, the adjective *homōnymos* ‘having the same name’ is already attested in Homer, and *polýsēmos* ‘having multiple meanings,’ in Democritus. Both adjectives were also borrowed into Latin as *homonymus* and *polysemos/polysemus* respectively. In English, *polysemous* has been attested since the beginning of the 19th century. Of the corresponding abstract nouns, *homonymy* goes back to ancient Greek *homōnymía*, while *polysemy* was coined as *polysémie* as late as 1887 by the French linguist Michel Bréal, whose use of the term however also subsumed *homonymy* (cf. Nerlich 2003: 58). Its antonym, *monosemy*, was introduced in the first half of the 20th century, also by French-writing scholars as it seems (Fr. *monosémie*).

Bréal introduced the concept “polysemy” in the context of semantic change, emphasizing the fact that the old sense does not disappear but, at least during some period of time, coexists with the new one. It is this process of “multiplication” of senses as well as the resultant state of one word having multiple senses which Bréal dubbed *polysémie*. Much of the subsequent discussion concerned the criteria for distinguishing polysemy from homonymy on the one side and from ambiguity or vagueness on the other.

The homonymy/polysemy problem has always been particularly pressing in lexicography (cf. Zöfgen 1989, Tarp 2009), where it has immediate practical consequences, due to the tradition of treating homonyms in separate entries but polysemes as a single entry with subentries. Fundamentally, two criteria have been proposed for distinguishing homonyms and polysemes. According to the etymological criterion, formally identical words are considered as homonymous if they go back to different etyma and as polysemous if they have the same etymon. The semantic criterion, on the contrary, is purely synchronic, stating that formally identical words are homonymous if speakers do not perceive any semantic relationship between them and polysemous, if they do perceive one. This latter criterion, which now prevails in linguistics, does not yield a neat separation but a continuum, since semantic relatedness is obviously a matter of degree. The reference to “speakers,” furthermore, opens up the possibility of divergent perceptions on the part of different speakers in one and the same case. In traditional lexicography and lexicology, a shift of syntactic category has also been viewed as a sufficient criterion for postulating homonyms, which is why the noun *hammer* and the verb *hammer* often have two separate entries in dictionaries despite their obvious semantic relationship. On that point, some modern scholars, especially cognitive linguists, depart from tradition, treating conversion as a kind of polysemy (cf. Dirven 1999).

The delimitation of polysemy and vagueness is even more elusive (cf. Cruse 1986: 49–83, Schneider 1988, Geeraerts 1993). Cruse argues that *cousin* in a sentence like *Sue is visiting her cousin* is not ambiguous between ‘male cousin’ and ‘female cousin,’ but simply vague; it only has one abstract sense covering both sexes and therefore constitutes a single ‘lexical unit’ in his terminology. *Bank* in *We finally reached the bank*, on the contrary, is said to be really ambiguous between the readings ‘financial institution’ and ‘side of river,’ which therefore constitute two different senses/lexical units. The variation within a single sense, as in *cousin*, he calls modulation, the variation of the

bank-type, contextual selection. While modulation is said to be “largely continuous and fluid in nature,” contextual selection of senses “proceeds in discrete jumps” (Cruse 1986: 52). Intuitively, this distinction might seem to be clear enough, but Cruse and Geeraerts show that the criteria proposed up to the present day for distinguishing vagueness and ambiguity—relations of meaning with other lexical units, within and across languages, zeugmatic effects, etc.—do not always yield clear results. A concept closely related to Cruse’s “modulation” is that of “dot object” in Pustejovsky (1995), which is intended to capture the subtle polysemy of, say, book in I liked the book (content) and Hand me the book (object).

The problem of polysemy is intimately tied to that of abstractness of meaning. Ever since the advent of structuralism a tendency can be observed to reduce the multiplicity of senses by postulating a more abstract core sense from which the concrete senses are said to be derivable by means of independently motivated rules of semantic extension or by tapping into contextual or encyclopedic information. Extreme manifestations of this theoretical position can be found in Ruhl (1989) or Pauly (2010). This latter author, for example, proposes to subsume all the senses of French *aller* ‘to go’ under one general “schematic form” characterized in the following terms: “ce verbe indique qu’étant donnés deux repères (spatiaux, temporels, subjectifs. . .), ces deux repères ne sont pas en rupture” (Pauly 2010: 302) [this verb indicates that, given two (spatial, temporal, subjective. . .) landmarks, these two landmarks are not at odds]. This general schematic form is said to “account for” uses of *aller* as in *je vais à Paris* ‘I go to Paris,’ *je vais y réfléchir* ‘I’ll think about it; lit. I go to think about it,’ or *je vais bien* ‘I’m fine; lit. I go well.’ Without going into details, one may say that this example nicely illustrates two problematic aspects of monosemism, viz. the vagueness of the definition of the supposed abstract core sense and the inexplicitness about how to derive exactly the observable senses from the postulated abstract sense. A third problematic aspect is constituted by overgeneration: The highly abstract schematic form quoted above is probably able to cover more verbs than just *aller* and more uses of *aller* than just those that are actually attested.

Concerning the relationship between the core sense and the secondary senses, a very useful conceptual distinction has been introduced in Cognitive Linguistics (cf. Lakoff 1987, Lewandowska-Tomaszczyk 2007). Cognitive linguists rightly insist that, though it is not possible in general to *predict* the specific senses from the core sense, these are nevertheless often *motivated* by the core sense. If you know that some language has a verb with the meaning “to go,” you cannot predict whether this verb can also be used to form the future tense, as is the case for French *aller*. However, the use of *aller* in the future periphrasis is clearly motivated by the core sense of this motion verb, at least upon reflection: If you displace yourself in order to perform some action, this action will inevitably take place somewhat later, since moving needs some time.

Accordingly, cognitive linguists do not think of the core sense as an abstract umbrella covering all the more specific uses, but as the starting point of a network of senses held together by relations of motivation, mostly of a metaphorical or metonymical nature (cf. already the graphic representation of such networks in Darmesteter 1887). If there is one central core sense surrounded by secondary senses which are all directly linked

to the core sense, it is customary to speak of a radial constellation (think of the spokes of a wheel: radial is related to Latin *radius* ‘spoke’). Latin *radius* itself, by the way, nicely illustrates what is meant by “radial constellation”: The core meaning was “stick,” which was then extended metaphorically in several directions, yielding “spoke,” “shuttle,” “(elongated kind of) olive,” “spur (of some birds),” or “penis.” Not all constellations, however, are of this kind. Sometimes, a specific sense in turn becomes the starting point for a further extension, and this game can in principle go on indefinitely, leading to quite intricate constellations, where the third sense need no longer show any direct semantic relationship to the first one. With a term borrowed from Ludwig Wittgenstein, the more complicated constellations of this kind are said to be held together by family resemblances. To take up our Latin example, one could possibly argue that the sense “ray” of *radius* was not due to a direct metaphor on “stick”, but rather one on the secondary sense “spoke”: In that case the sun and its rays would be likened to the hub and its spokes.

As I said at the beginning, the discussion of polysemy/homonymy began in the realm of rhetoric and philosophy, then moved to (historical) semantics and eventually settled down in lexicology/lexicography. In more recent times, computational linguists and psycholinguists have also become interested in the phenomenon, contributing valuable insights that can no longer be by-passed by linguists. From the psycholinguistic literature, I would particularly like to draw attention to Murphy (2007). This author explicitly argues that the mental lexicon is not parsimonious, because its main purpose is “to allow fast and accurate language use” (2007: 60), not to minimize storage (a typical obsession of computer scientists, but also of many structuralists and generative linguists). According to his theory, “whether a sense will be mentally represented [. . .] is largely determined by whether the sense has been experienced before, and how often” (2007: 65; cf. also de Vaan et al. 2007). Murphy concedes that people do generalize over similar patterns of polysemy, but such generalizations or rules are said to be activated only for the interpretation of novel uses of words. Both the “animal” and the “meat” sense of *chicken*, for example, are stored separately in the mental lexicon, as suggested by the fact that the use of one of the senses slows down the subsequent use of the other sense (Klein and Murphy 2001). At the same time, the generalization over similar polysemes allows the speaker to interpret a novel use like *I ate gnu for the first time during a safari in Kenya*. Importantly, however, they do not use such generalizations in order to minimize storage, listing only the “animal” sense of *chicken* and deriving the “meat” sense by rule. Another psycholinguistic study of considerable interest in our context is Brown (2008). Participants were asked to judge as quickly and accurately as possible whether a phrase presented on a screen “made sense” or not. The test items contained verb senses with four different degrees of semantic relatedness: same sense (e.g. *cleaned the shirt/the cup*), closely related sense (e.g. *broke the glass/the radio*), distantly related sense (e.g. *ran the track/the shop*), unrelated sense (e.g. *banked the plane/the money*). When moving from one phrase to another, the quickness of the answers was correlated positively with semantic closeness. The smooth progression observed even between distantly related and unrelated senses is interpreted by the author as evidence in favor of a single-entry theory where senses constitute hierarchically ordered subentries or

subentries of subentries, according to the semantic distance from the central meaning, and where homonyms are also accommodated at the periphery of the single entry. In the face of the MEG study by Pykkänen et al. (2006), however, the author does not exclude the possibility that homonyms might have to be granted a separate entry after all.

19.3 THE MONOSEMY–POLYSEMY DEBATE IN WORD FORMATION

The temptation to make the multiplicity of senses disappear by having recourse to abstract meaning is even greater in the analysis of patterns of word formation than in that of simplex words. It has been lurking around from the very beginnings of semantics as a scientific discipline (cf. Bréal 1868) and its appeal is unlikely to abate soon.

19.3.1 The Polysemy of the English Suffix *-er*

The English suffix *-er* has been at center stage in discussions on polysemy in word formation for almost three decades now. A review of this abundant, and relatively well-known, literature is beyond the scope of a handbook chapter (cf. Lieber (2004: 17) for references). I will therefore limit myself here to a brief discussion of Plag's (2003: 89) treatment, which serves my purpose well by being both short and illustrative of a style of description that thrives for unitary meanings at any cost:

Apart from performers of actions we find instrument nouns such as blender, mixer, steamer, toaster, and nouns denoting entities associated with an activity such as diner, lounge, trainer, winner (in the sense “winning shot”). Furthermore, *-er* is used to create person nouns indicating place of origin or residence (e.g. Londoner, New Yorker, Highlander, New Englander). This heterogeneity suggests that the semantics of *-er* should be described as rather underspecified, simply meaning something like “person or thing having to do with X.” The more specific interpretation of individual formations would then follow from an interaction of the meanings of base and suffix and further inferences on the basis of world knowledge. (Plag 2003: 89)

Why is such an abstract (or “underspecified”) analysis of the semantics of *-er* potentially problematic? *Mutatis mutandis*, the objections raised against the abstract unitary analysis of *aller* in Section 19.2 also apply to the case at hand, at least partially. The abstract meaning proposed by Plag (“person or thing having to do with X”), it is true, escapes the reproach of being unduly vague, but it is not really a unitary meaning, since the formula employed contains a disjunction (“person *or* thing”). This shortcoming, however, could be remedied by substituting a still more abstract noun such as “entity” to the disjunction “person or thing.” The second objection concerns the indeterminacy of the mechanisms

made responsible for specifying the concrete senses on the basis of the abstract sense. According to Plag, these are the “interaction of the meanings of base and suffix” as well as “inferences on the basis of world knowledge” (Plag 2003: 89). In order to assess this claim, the cognitive-linguistic distinction between prediction and motivation may be helpful. While it is doubtlessly true that the concrete meanings of the examples given by the author can be said to be motivated in a straightforward way by their bases, it is much less obvious that it is really possible to predict the exact range of attested or acceptable meanings by taking into account exclusively the meaning of base and affix as well as general encyclopedic knowledge. The major shortcoming of abstract analyses of this kind is massive overgeneration. The suffix, though highly versatile, is not used, for example, to denote a “dialect spoken in X” (e.g. [*I speak the*] **Bostoner*, **Sussexer*), nor a “tree bearing fruit X” (e.g. **peacher*), nor a “follower of X” (e.g. **Hitlerer*), to give but a few representative examples. A comparison with semantically related suffixes of other languages shows that such restrictions are a conventional feature of English. Finnish, for example, uses the same suffix *-lainen* to refer to inhabitant names (e.g. *Helsinki* → *helsinki-läinen*) and followers (e.g. *Marx* → *marxilainen*). Such language-specific regularities are difficult to accommodate in an analysis where the only conventional meaning is “entity having to do with X.”

A way out for defendants of abstract meanings in cases like these could be to claim that *-er* is not really impossible but only sounds odd because it is blocked by a more specialized rival pattern. Now, while well-entrenched rival patterns or words undoubtedly account for some of the otherwise unexplainable lacunae, blocking, which by the way is hardly ever stated explicitly, certainly is not a fix-it-all device for saving over-abstract analyses. Without assessing the merits of a blocking account in cases like *Hitlerite*/**Hitlerer*, *Marxist*/**Marxer*, or *peach tree*/**peacher*, it seems quite obvious, for example, that such an account must fail to explain why **Bostoner*, **Sussexer*, etc., in the sense of “dialect spoken in place X” are odd, since there is no rival pattern. It is quite imaginable that *-er* could one day extend its reach to cover even referential categories such as those mentioned above, but that should then be viewed as an instance of semantic extension, of language change, rather than simply as a banal realization of possible words.

19.3.2 The Semantic Interaction of Base and Affix

One issue which deserves closer scrutiny is the mechanism characterized by Plag in Section 19.3.1 as “interaction of the meanings of base and suffix.” One of the earliest scholars who explicitly proposed to use it in combination with abstract meanings for the purpose of bringing order to the chaos of suffixation was Gamillscheg (1921: 22–5). He had discovered an almost perfect correlation between the distribution of the instrumental and the locative sense of the Latin suffixes *-aculum* and *-torium* (as well as its French descendant *-oir*) and the valency of the verbal base or the derivatives: Transitive verbs give rise to instrument nouns (e.g. *gubernaculum* ‘helm’ ← *gubernare* ‘to steer’),

intransitive verbs to locative nouns (e.g. *cenaculum* ‘dining room’ ← *cenare* ‘to dine’). He proposed to account for this correlation by assigning to the suffixes in question the meaning of a pure action noun—a relatively marginal sense of the suffix in Latin, by the way—and to derive the concrete senses by having recourse to what he thought to be “privileged associations”: The concept “steering,” he maintained, was most strongly associated with the instrument used for steering (helm), while the concept “eating” was more strongly associated with the place where the action takes place (dining room) than with the instruments used (fork, knife, spoon, plate, etc.) or the meal. Gamillscheg did not provide independent evidence for the postulated strength of his “associations,” which was probably more intuitively evident to him than to me. But we should not throw stones while living in glass houses. Many arguments in contemporary work on word formation based on “inferences,” “pragmatics,” “encyclopedia,” or the like do not fare much better as regards their empirical underpinning.

It is common, for example, to state that the suffix *-ist* simply means “person having to do with X,” while the rest is said to follow straightforwardly from encyclopedic knowledge provided by the base. Such an account seems to nicely explain why a flutist *plays* the flute, while an Arabist *studies* Arabic. However, isn’t speaking much more closely associated to language than studying? But why then aren’t all speakers of Arabic Arabists, rather than only those who “normally from outside the Arab World specialize in the study of the Arabic language and culture”? And why can’t the Romans be called Latinists? It seems that somehow we cannot avoid stating explicitly in our description of this suffix that—in the present state of English—with names of languages the “specialist” sense is conventionalized, while the “speaker” sense, which we should expect to exist if the meaning of *-ist* were really “person having to do with X,” is at best a remote possibility, a potential future extension of the suffix’s use. The “specialist” sense of *-ist* is clearly contextually determined, but nevertheless conventional and not the result of automatic inferences. As this example suggests, arguments based on the interaction of base and affix deserve much care if we want to avoid the trap of pseudo-explanations.

19.3.3 Structuralist Polysemophobia

The tendency to postulate abstract unitary meanings has been taken to an extreme by Émile Benveniste and his numerous followers in the structuralist tradition (cf., among many others, Serbat 1975, Laca 1986, Amador Rodríguez 2009). While most authors who practise this style of analysis do not feel compelled to justify its theoretical implications, Benveniste (1948) apodictically states at the beginning of his classic:

Our demonstration will verify a simple principle: If two living formations are in competition, they cannot have the same value. And vice versa: Different functions associated with one and the same form must have a common base. It is the task of linguists to find these values, generally barely visible and often deeply hidden. (Benveniste 1948: 6 [translation, F. R.])

Applied to suffixation, the topic of Benveniste's book, this means that: (a) if a suffix at first sight has several senses, closer scrutiny will always allow finding a common abstract denominator; (b) if a particular meaning is expressed by more than two at first sight synonymous suffixes, closer scrutiny will always allow the discovery of some systematic difference. Note that Benveniste did not mean the sentence quoted above as a heuristic principle, as a morphological version of Ockham's razor (as such, it would of course be quite innocent and acceptable), but as a factual statement about word-formation systems.

Now, it is easy to see that both the No-Polysemy and the No-Synonymy part of Benveniste's statement are wrong. The existence of perfectly synonymous patterns of word formation is beyond doubt: Spanish, for example, has several dozen suffixes for forming ethnic adjectives, for example, *toled-ano* 'from Toledo,' *madril-eño* 'from Madrid,' *granad-ino* 'from Granada,' *cartage-ero* 'from Cartagena,' *barcelon-és* 'from Barcelona,' etc. The choice of suffix is mostly lexically conditioned, though some phonological and geographical factors also come into play. Semantically, however, all suffixes are perfectly equivalent. Examples apt to falsify the No-Polysemy part are not difficult to find either. Just think of the two uses of the suffix *-itis* in English and other European languages. In its literal use in medicine, this suffix designates diseases (more specifically, inflammations): *appendicitis* (← *appendix*), *pancreatitis* (← *pancreas*), *tonsillitis* (← *tonsil*), etc. Beside this original use, the suffix has also been used—since around 1900 in English—to refer to figurative “diseases” in jocular language: *telephonitis* ‘excessive fondness for telephoning,’ *Wagneritis* ‘excessive fondness for Wagner,’ etc. The metaphorical link between the first and the second series is immediately obvious to any competent speaker of English, that is, the figurative use is clearly motivated by the medical use. A reasonable account of this state of affairs would seem to be to say that English has two separate suffixes *-itis* linked by a metaphorical relation. A Benvenistian monosemist, on the contrary, would have to claim that with the advent of the figurative use the suffix *-itis* acquired a more abstract meaning covering both real diseases (inflammations) and the peculiar states of obsession designated by the figurative series. It is not easy to see how such an abstract meaning could be defined. The most plausible formula that seems to suggest itself is “disease related to N,” but this solution would have two drawbacks. On the one hand, one could argue that the formula does not really represent a unitary meaning, since it exploits the polysemy of the noun *disease*, which in English designates both real diseases in the medical sense of the word and figurative “diseases,” as in *the disease of materialism*. On the other hand, defining *-itis* as “disease” instead of the more concrete “inflammation” would not account correctly for the limitations of use of this suffix in the language of medicine, where not just any disease is compatible with this suffix: While *arthritis* designates an inflammation of a joint, it cannot, for example, refer to the degeneration of a joint, a type of disease which requires the suffix *-osis* (*arthrosis*). Similarly, addictions have a preference for *-ism* (cf. *alcoholism*, *cocainism*, *heroinism*, etc.) and are incompatible with *-itis* in the relevant sense (**alcoholitis*, etc.). As one can see, defining *-itis* as “disease” leads to massive overgeneration.

19.3.4 A Set of Criteria for Unitary Meanings

More in general, let us now ask under what conditions proposing a unitary meaning can be considered well-founded. Kleiber (2010: 137–43) has recently proposed a series of criteria for simplex words, two of which at least can also be applied to word formation. A first criterion requires patterns to designate homogeneous categories. As I have argued in the last paragraph, this requirement is not met in the case of our two series of words in *-itis*: Inflammations and figurative diseases do no more constitute a homogeneous category than the common house mouse and the computer mouse. The second criterion requires that unitary meanings not only cover all attested semantic values, but at the same time exclude impossible ones. Again, we have seen that an abstract meaning for *-itis* such as ‘disease’ leads to undesirable overgeneration. The heterogeneity of the two suffixes is also underlined by the fact that their bases constitute complementary sets. Consequently, there is no attested formation in *-itis* that pertains to the literal and the figurative set at the same time, which is why two other criteria put forward by Kleiber are inapplicable. One of these requires that the assignment of truth values must apply to all uses of a supposedly unitary word, while the other one restricts the possibility of thinking of all semantic values simultaneously to words with a unitary meaning.

19.3.5 Making Abstract Meanings More Articulate

Many linguists will probably be prepared to accept that we cannot get very far in the study of the semantics of word formation with formulae of the type “X that has to do with Y.” One way out that has been proposed in the literature is to make the abstract formulae themselves more articulate.

Plank (1981: 46–8) proposed a “holistic” theory of meaning in word formation where affixes are tied to complex conceptual schemata containing variables and where speakers have some freedom to match the base with one of the variables (a process called “thematization” by him). The German separable verb-forming prefix *auf-* ‘up’, for example, is said to have the following complex meaning: “to apply something (PATIENT) by way of an appropriate action (MODE) or appropriate means (MEANS or INSTRUMENT) to the surface of something (PLACE).” According to Plank, such an articulate schema is able at the same time to provide a unitary account for many semantic effects of the prefix and the variability of the syntactic category of the base:

- (1) a. MODE (verbal base): *aufnähen* ‘to sew on’ (← *nähen* ‘to sew’);
- b. MEANS/INSTRUMENT (verbal or nominal base): *aufkleben* ‘to stick on’ (← *kleben* ‘to stick’), *aufspießen* ‘to spear’ (← *Spieß* ‘spear’);
- c. PLACE (nominal base): *aufbahnen* ‘to lay out (a corpse)’ (← *Bahre* ‘bier’);
- d. PATIENT (nominal base): *aufbürden* ‘to load sth. on to sb.’ (← *Bürde* ‘burden’).

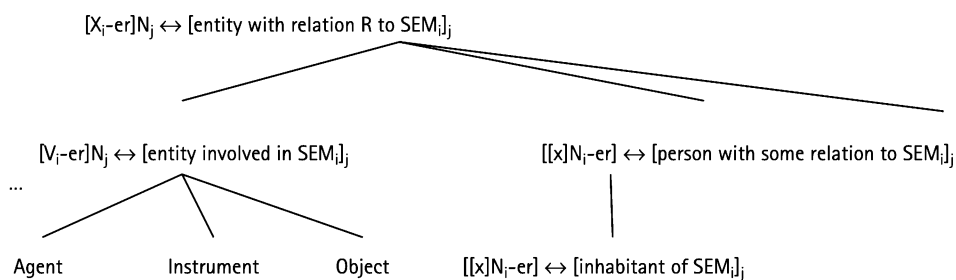
The variables not bound by the base of the derived verb are said to be realized in the syntax according to general linking rules. As far as I can see, this approach, which is only sketched in Plank's book, has never been worked out in detail either by the author himself or by anybody else.

Another approach based on elaborate schemata has been worked out more explicitly in Lieber (2004). Her approach is couched in a framework of decompositional semantics. (Contrary to Plank's it is also compositional, i.e. the meaning of a complex word is assembled bottom-up, not top-down as in a holistic approach. Affixes are treated as signs with separate lexical entries.) Lexemes and affixes are thought to consist of a "skeleton" containing only the most abstract semantic features and a "body" providing more concrete encyclopedic information, if there is any. The conceptual combination of bases (in compounding) or base and affix (in derivation) is effectuated by a general principle of "co-indexation"—the analogue of Plank's "thematization"—which, roughly speaking, co-indexes the highest non-head argument with the highest head argument. The meaning of *truck driver*, for example (cf. Lieber 2004: 55), is derived in the following two steps: In a first step, the external argument of the verb *drive*, which is higher on the hierarchy than the internal argument, is co-indexed with the only argument of the suffix *-er* (the referential argument); hence, *driver* comes to mean "person driving N" and does not refer to the vehicle driven. In a second step, the internal argument of *drive* that remained free is co-indexed with the referential argument of *truck*, which thereby takes over the theme role in semantic interpretation ("person that drives a truck"). Now, let us look at the way in which we get the reading of *city employee*. Since the suffix *-ee* preferentially binds the internal argument of its verbal base, *employee* comes to mean "person that N employs". This time it is the external argument that remains free. As a consequence, the noun *city* is co-indexed with this external argument, which yields the correct reading "person that the city employs". As the reader may suspect, things do not always work out that neatly. The book itself acknowledges some of the limitations of this theory and proposes additional mechanisms to overcome at least some of them. One such mechanism is "paradigmatic extension" (Lieber 2004: 72–5). In order to explain why English *-er* does not always bind the external argument (cf. *keeper* 'thing worth keeping,' *walker* 'a frame device used to support an infant learning to walk,' *diner* 'place for dining,' etc.), Lieber assumes that affixes can go beyond their original meaning in case the word-formation system does not provide a dedicated pattern for the meaning to be expressed. One cannot exclude that a "pull effect" of this kind exists, but it is certainly neither systematic nor automatic, and should probably better be viewed as a factor occasionally operative in semantic change rather than as part of a synchronic theory of word formation.

19.3.6 Hierarchically Ordered Constructional Schemas and Radial Categories

Another way of accounting for polysemic word-formation patterns in a synchronic grammar has recently been proposed in the framework of Construction Morphology.

Booij (2010: 76–84) argues that different, but semantically related uses of a formal pattern should be represented in the form of hierarchically ordered constructional schemas such as the one given below. Agent, instrument, and object schemas, which head lexemes that are the Dutch equivalents of English rider, poker, and keeper, are in turn headed by a schema with the more abstract meaning “entity involved in SEM_i” (where SEM_i represents the meaning of the verbal base), and this schema is itself headed by a still more abstract one with the meaning “entity with relation R to SEM_i” (where SEM_i represents the meaning of bases of any syntactic category). The main difference with respect to Plag’s treatment in Section 19.3.1 is that “subschemas for the different interpretations such as Agent, Instrument, Object, etc. . . . function as the level at which new. . . nouns are coined” (Booij 2010: 79). The hierarchy above them “specifies. . . their common properties,” but “does not give direct expression to the nature of the semantic relationships” (2010: 80), which seem to be considered of diachronic relevance only.



This treatment has the advantage of acknowledging the central importance of low-level schemas for coining (and interpreting, one might add) new complex words. Booij also shows that this level of representation is important for stating paradigmatic relationships in the lexicon. In Dutch, for example, inhabitant names in *-er* form their female counterpart with the suffix *-se* (e.g. *Amsterdammer* → *Amsterdamse*), while other denominal *-er* nouns select *-ster* (e.g. *wetenschapper* “scientist” → *wetenschapster*). However, Booij’s account is not exempt from potentially problematic implications. First, it is not clear to me whether the notion of “default inheritance,” which is central to the hierarchical layering of schemas in Construction Morphology, does not also lead to overgeneration. It is true that a lower pattern takes precedence if it has another specification for some property, but what about all those cases where no conventionalized low-level patterns exist, but which are nevertheless unacceptable? A second potential problem concerns the absence of a direct expression of semantic relationships between schemas. As we have seen in the treatment of the suffix *-itis*, the metaphorical relationship between the medical and the jocular use of this suffix is vital even from a synchronic point of view and should therefore be given explicit expression in the grammar.

Both of these problems are avoided if polysemy is treated as a radial or family-resemblance constellation. In that case, all patterns are concrete in the sense that they define a set of possible words. (The formation of a particular complex word can only be hindered

by blocking, either by a well-entrenched lexeme or a synonymous rule.) To the extent that the semantic relationship between two patterns is still perceived synchronically, it is explicitly stated in the description as a relationship of motivation. If no synchronically relevant relationship exists, the patterns are treated as homonymous, even if from a historical point of view one of the patterns was derived from the other through sense extension. What such a treatment would miss in the case of *-er* is the generalization that all nouns refer to an entity, but it is unclear a priori that this is really a grammatically relevant generalization. Construction morphologists should try to gather more substantial empirical evidence on this point.

19.4 CROSS-LINGUISTICALLY FREQUENT POLYSEMIC CONSTELLATIONS

In this section I would like to briefly enumerate some polysemic constellations that have been observed very often cross-linguistically in word formation. Though semantic change can sometimes take a quite unexpected course, many extensions can be observed over and over again (cf. Apresjan's (1974) concept of "regular polysemy"). From a diachronic point of view, some of the examples mentioned in this section constitute cases of grammaticalization (Heine et al. 1991, Hopper and Traugott 2003).

The concretization of abstract nouns has already been identified as a universal phenomenon by Paul (1880). Action nouns often at the same time refer to one of the participants (agent, theme, instrument) or to the setting (place, time) of the action, as well as to the result (resultant state, effected object). On the rich history of research in this domain, cf. Rainer (1996), for a recent contribution with abundant references to the most recent literature, Melloni (2011). Due to their simpler semantic structure, quality nouns, when concretized, normally denote the bearer of the quality (e.g. *a beauty*). Status nouns also tend to be highly polysemous, denoting for example the time in office (e.g. *during his popeship*), the corresponding territory (e.g. *dukedom* 'duchy') or the collectivity of persons involved (e.g. *the peasantry*).

About half of the languages of the world that build agent nouns morphologically also use them as instrument nouns (cf. Luschützky and Rainer 2011). The percentage is even slightly higher with respect to word-formation patterns which serve at the same time for the formation of instrument and place nouns (Luschützky and Rainer 2013), such as *-aculum* in Latin (cf. *gubernaculum* 'helm,' *cenaculum* 'dining room'). Contrary to what we have seen with abstract nouns, there is little evidence that these two cases of syncretism are due to direct semantic extension, as we will see in Section 19.5. Both instrument and place nouns, however, often have their origin in action nouns, as mentioned in the last paragraph.

Evaluative patterns (cf. Jurafsky 1996) tend to be accompanied by value judgments (positive or negative, depending on the base noun), which often give rise to separate

meliorative or pejorative patterns in diachrony. The broad diffusion of this polysemic constellation is transparently attributable to pragmatic inferences: Small things are either felt to be cute (e.g. *doggie*) or despised (e.g. *kinglet*), big things either despised (e.g. Italian *donnone* ‘big woman’) or, on the contrary, admired (Peninsular Spanish *cochazo* ‘big, great car’).

Polysemy is less diverse with adjectives, whose meanings are generally less “rich.” Material adjectives (e.g. *golden cup*) and ornative adjectives (e.g. *oily trousers*) frequently develop a resemblance meaning (e.g. *golden sand*, *oily character*). Adjectives of resemblance themselves often seem to give rise to an approximative sense (e.g. *brutish/reddish*). Concerning intensifying affixes, we can often observe that attenuation is expressed with diminutive suffixes, while many boosters have their origin in locative affixes expressing high location/rank (cf. *hyper-*, *super-*, *archi-*, etc.) or transgression of a limit (cf. *trans-*, *extra-*, *ultra-*, etc.). Relational adjectives are generally said to be able to express any relation that can plausibly be established between the base noun and the head noun. To the extent that this is true, the semantic variation (cf. *musical instrument* vs. *musical clock* vs. *musical instruction*, etc.) would have to be considered as induced by the context on the basis of encyclopedic information.

In the domain of verbs at least three recurrent polysemic constellations should be mentioned. First, denominal verbs are also often thought to have only a minimal semantics, leaving the trouble of specifying the exact meaning to context and/or encyclopedia (cf. Aronoff 1980). There is certainly some truth to this commonplace, since the same relations show up with the same concepts in language after language, but nevertheless we also find striking differences (Hungarian, but not English or German, for example, is free to use denominal verbs to express the concept “to drink N”). Apparently, the pragmatic-encyclopedic processes are heavily constrained by language-specific conventionalization. A second frequent constellation is what is referred to as “verbal plurality” (Dressler 1968) or “pluractionality” (Wood 2007): It is often the case that one and the same pattern of word formation—often reduplication—adds features of quantification, distribution, repetition, duration, or diminution—or some subset thereof—to the verbal base. A third frequent polysemic constellation can be found with locative prefixes or particles, which are particularly prone to develop *aktionsart* or aspectual meanings (cf. *up* in *to eat up the soup*). Locative affixes also frequently show a corresponding temporal meaning, but this widespread metaphor is not limited to the verbal domain (cf. *prealpine* vs. *pre-Elisabethan*).

19.5 POLYSEMY IN DIACHRONY

In work on semantic change, complex words are generally treated on a par with simplex words, apparently because of a widely-held belief that both behave in the same way. Jaberg (1905) even explicitly claimed that semantic change in affixation was

always the result of semantic change in individual words plus reanalysis. In Rainer (2005a) I tried to show that, on the contrary, patterns of word formation can also directly undergo semantic change, normally by being combined with metaphor or metonymy at the moment of creating a neologism. We have already encountered one case in point when discussing the suffix *-itis*. The figurative use of this suffix came about by combining the original pattern with a metaphor: Instead of “real disease related to N” it was taken to mean “figurative disease related to N.” The stylistic effect of the figurative use is precisely due to the implicit recategorization of some states of excessive fondness as diseases. Note that at no moment in history did one of the medical terms in *-itis* change its meaning in a way that might have given rise to a reanalysis of the suffix, as Jaberg’s hypothesis would lead one to expect. As we have just seen, metaphor implies a shift in semantic category when applied to a pattern as a whole. However, the metaphor or metonymy can also apply to the base word only, leaving the semantic category of the output intact. Spanish, for example, has a suffix, *-uno*, originally forming relational or similitive adjectives only on the basis of animal nouns (e.g. *lobuno* ‘relating/similar to wolves’ ← *lobo* ‘wolf’). During the Renaissance, however, speakers also began applying this suffix to human bases, thereby implicitly assimilating the persons denoted to animals (e.g. *lacayuno* ‘relating/similar to servants’ ← *lacayo* ‘servant’). The repeated application of such metaphorical and metonymic extensions can lead over time to quite extended radial or family-resemblance structures in synchrony. Heine et al. (1991: 78–97), for example, show how in Ewe the noun *vi* ‘meaning “child” developed into a suffix with an array of meanings linked to the original meaning through ‘channels of conceptual expansion’ such as ‘young’ > ‘inexperienced’ > ‘that has not yet passed an exam’ > ‘unsuccessful’, ‘small’ > ‘insignificant’, ‘small’ > ‘delineated part of a mass’, ‘member (of a family)’ > ‘member (of a community)’. Interestingly, as the authors observe, “some of the meanings found in European and other languages, such as those expressing affection of tenderness. . . , appear to be absent in Ewe” (Heine et al 1991: 86–7). This again proves that it has to be stated in the grammars of single languages which sense extensions have actually taken place. Another detailed discussion of such a fragmentation process is provided by Rainer (2003), who shows how the Spanish suffix *-azo*, from the original meaning ‘blow dealt with N’ (e.g. *porrazo* ‘blow dealt with a club’ ← *porra* ‘club’), developed, among others, the senses ‘shot fired with N’ > ‘noise produced by a shot with N’ > ‘sharp sound produced with N’ (e.g. a trumpet) > ‘sharp cry produced by N’, ‘blow dealt on bodypart N’, ‘wound caused with N’, ‘coup staged by N’ > ‘spectacular political measure taken by N’. The fragmentation process, of course, need not necessarily rely on metaphorical and metonymic extension; cases of reanalysis or borrowed words, for example, can also constitute the starting point for new sub-patterns (cf. Mühleisen 2010 on the development of the English suffix *-ee*). This kind of continuous process of semantic fragmentation of patterns, which can be observed over and over in word formation, is in itself a strong caveat against monosemism.

Another diachronic topic worth discussing is the relation between polysemy and homonymy. One commonplace of studies in word formation is that the frequent

expression of the concept of agent and instrument with one and the same pattern is due to an extension of the agent reading to an instrument reading as a consequence of a supposed close semantic affinity between these two conceptual categories. English *-er* (cf. *speaker* vs. *cooker*) as well as French *-eur* (cf. *mangeur* ‘eater’ vs. *compteur* ‘meter’ ← *compter* ‘to count’) are frequently cited examples. This account seems to have such an irresistible intuitive appeal that it has hardly ever been challenged. However, in a detailed diachronic reconstruction of the history of French *-eur* and its Romance cognates (cf. Rainer 2011), I have shown that, contrary to received opinion, no semantic extension has ever been involved in the evolution of the exclusively agentive Latin suffix *-tor*. In Catalan and Occitan as well as some French and Italian dialects, Latin *-tor* merged with instrumental-locative *-torium* for purely phonological reasons. In these varieties, the resulting suffix consequently shows a threefold polysemy (or homonymy, for that matter): agentive-instrumental-locative. This threefold polysemy/homonymy has then been transferred via loan words to Spanish and some other varieties where *-tor* and *-torium* remained distinct (cf. Sp. *cazador* ‘hunter’ vs. *destornillador* ‘screwdriver’ vs. *comedor* ‘dining room’). In Standard French and Italian, which have also kept the two Latin suffixes distinct, the instrumental pattern has been established essentially since the second half of the 18th century, mainly through loan translations from English (cf. E. *ventilator* > F. *ventilateur*, I. *ventilatore*, etc.) or ellipsis of the head noun of lexical noun phrases denoting instruments, of the type I. *apparecchio misuratore* ‘measuring instrument’ > *misuratore* ‘id.’ Significantly, these two languages do not have the locative sense. A cross-linguistic study going beyond Romance and Indo-European on the agent-instrument “polysemy” (cf. Luschützky and Rainer 2011) yielded similar results, as well as a follow-up study on the instrument-place “polysemy” (cf. Luschützky and Rainer 2013). The main factors involved were again borrowing, phonological merger, ellipsis and, above all, the concretization of action nouns. No evidence for a direct sense extension from agent to instrument or instrument to place has been found. While the existence of polysemous patterns in word formation in general is beyond doubt, as several examples adduced in this chapter have illustrated, one could certainly make a case for treating English *-er*, Spanish *-dor*, Latin *-aculum*, and similar suffixes as homonymous rather than polysemous, at least in some of their uses. Students of word formation still have to dress a reliable catalogue of criteria that allow distinguishing in synchrony real polysemy from multifunctionality which is only the accidental result of factors such as phonological merger, borrowing, or ellipsis.

19.6 CONCLUSION

My impression is that the in-depth study of polysemy in word formation has just begun. Much remains to be done, both in synchrony and in diachrony. Detailed comparisons of closely related languages would be particularly helpful. Many of those who have

dedicated their time and effort to this area have done so with the intention to make polysemy disappear by portraying it as an automatic consequence of more abstract unitary meanings. As good disciples of William Ockham we should of course always avoid postulating unnecessary entities, but psycholinguistic studies as well as unprejudiced observation show that speakers cope well with massive polysemy and homonymy. Apparently it is preferable for them to have many low-level patterns that can be immediately applied in comprehension and production. These patterns may of course form complex constellations held together by metaphor and metonymy or show a relationship of subsumption under more abstract patterns. Even in this latter case, however, we need the low-level patterns in order to be able to account for all kinds of distributional restrictions as well as innumerable subtle differences observable among closely related languages.

CHAPTER 20

DERIVATIONAL PARADIGMS

PAVOL ŠTEKAUER

20.1 INTRODUCTION

AN inherent part of almost all discussions of the differences between inflection and derivation used to be the claim that inflectional morphology is paradigmatic while derivational morphology is not. In fact, as admitted by van Marle (1985: 14), “the relationships between elements in *absentia* are not a ‘popular’ field of investigation” not only in the generative rule-based tradition of word formation. Plank (1994: 1672), for example, maintains that inflectional categories form “a more or less closed system, being alternatively realized by a limited number of terms,” while derivational categories are “not part of any well-organized categorial system,” and Dressler (1989: 8) assumes that inflectional morphology “is typically organized in paradigms. . . whereas the paradigmatic organization of DM [derivational morphology]. . . is much weaker.”

Parallel to the traditional belief of irregularity and unpredictability of derivational morphology and, by implication, of the peripheral position of derivational paradigms, there are views postulating the paradigmatic nature of word-formation in accordance with Stump’s (1991: 710) assumption that “there can be no principled objection to the notion of a derivational paradigm. . .” and that “many of the arguments that motivate the postulation of paradigms in the inflectional domain have straightforward analogues in the domain of derivation” (Stump 2005: 65). From a different perspective, van Marle (1994: 2929) does not accept the view that “paradigmatics is of little or no importance to derivational morphology. The only conclusion which is justified is that in derivation paradigmatic structure may manifest itself in a fundamentally different way from the way it does in inflection.”

This chapter therefore aims to map different views of derivational paradigms and their nature. A brief introduction to the concept of *paradigm* (Section 20.2) is followed by pointing out the characteristic features of inflectional paradigms (Section 20.3). Section 20.4 discusses basic similarities and differences between inflectional and derivational paradigms against the fuzzy nature of the inflection—derivation opposition in general.

Section 20.5 presents various approaches to derivational paradigms and Section 20.6 identifies what seems to be the common basis of the concept of derivational paradigm.

20.2 ON THE NOTION OF PARADIGM

The term paradigm has been used in linguistics in several different meanings. Their scope (and diversity) follow from Saussure's broad concept of *associative relations* (1959: 125–6). To use Saussure's examples, one type of paradigmatic relations is represented by French lexemes like *enseignement* 'teaching,' *enseigner* 'teach,' *enseignons* '(we) teach,' etc., where one element, the radical, is common to all the related units. The paradigmatic organization may also rest on a common affix, for example, *enseignement* 'education,' *armement* 'armament,' *changement* 'change,' etc. Furthermore, the organizing element may be a common concept, as Saussure assumes, and illustrates this idea with *enseignement* 'teaching,' *instruction* 'instruction,' *apprentissage* 'learning,' *education* 'education,' etc. Paradigmatic relations may also refer to different types of organization of lexical units in the mental lexicon (synonymy, homonymy, polysemy, hyponymy, hyperonymy, etc.); and last but not least, they may refer to classes of lexical units that can assume the same slot in the syntagmatic organization of human speech. The most typical associative groupings in Saussure's view is the *inflectional paradigm* organized around a common base. In fact, this is how morphologists tend to conceive of this notion. Let us therefore first have a look at this common understanding of paradigm.

20.3 INFLECTIONAL PARADIGM

There are two basic notions of inflectional paradigm: first, it is a *general pattern*, a system of slots for relevant morphosyntactic categories each of which is realized by a specific formal marker; second, it is a *specific realization* of the general pattern in individual lexemes. For illustration, in Slovak, the declension of nouns is determined by formal gender (masculine, feminine, neuter), case (nominative, genitive, dative, accusative, vocative, locative, instrumental), and number (singular, plural). Thus, each slot represents a cumulation of three grammatical meanings represented by a single inflectional morpheme.

This principle is reflected in Pounder's (2000: 85) distinction between two concepts of inflectional paradigm:

- (a) Paradigm as a pattern valid for a lexico-syntactic class or subclass, i.e. a general morphological structure represented as a set of available paths between a base and appropriate operations;

- (b) Paradigm as a set of specific forms corresponding to one lexeme, i.e., an individual or “lexical” paradigm defined as a structured pattern of “instructions” for operations on stems, where feasible.

This idea is captured by Stump when he says that “inflectional rules apply to realize a pre-existing set of morphosyntactic feature specifications” (1991: 681).

A general feature of the inflectional paradigm is assumed to be a *common stem*. In addition, van Marle (1994: 2929) specifies the following characteristics: the individual members of a paradigm share prominent features (e.g. Latin *laudare*: *laudo*, *laudas*, *laudat*, *laudamus*, *laudatis*, *laudant* share not only the same base but also “nonperfect,” “indicative,” “active,” “present,” and “first conjugation”); the features in terms of which they differ are relatively marginal (person and number in the above example); paradigms are syntax-oriented in character; forms may be generated in an “automatic” way; a morphological operation never results in a change of word-class.

The last point is disputable as evidenced by cross-linguistic data showing that inflection may sometimes take over the function of derivation, that is, the coining of new words, accompanied by the change of word-class. Sapir (1965: 57), for example, provides a few instances of inflectional markers in Kujamaat Joola functioning as derivational affixes. For illustration, the passive suffix *-i* combines with *wɔl* ‘give birth’ to form *kawɔli* ‘birth_N’. Fairly widespread counter-evidence is provided by the process of formation of a new word by change of inflectional paradigm. This phenomenon, traditionally called “conversion” or “zero derivation,” has been labeled by various names emphasizing its inflectional nature. So, for example, Waszakowa (1993) and Szymanek (2010) speak about paradigmatic derivation, and Dokulil (1962) about transflexion. The following example comes from Szymanek (2010: 234):

(1)	Adjective ZŁY ‘bad, evil’			Noun ZŁO ‘badness, evil’
	<i>masculine</i>	<i>feminine</i>	<i>neuter</i>	
nominative	zł-y	zł-a	zł-e	zł-o
genitive	zł-ego	zł-ej	zł-ego	zł-a
dative	zł-emu	zł-ej	zł-emu	zł-u
accusative	zł-y/ego	zł-ą	zł-e	zł-o
instrumental	zł-ym	zł-ą	zł-ym	zł-em
locative	zł-ym	zł-ej	zł-ym	zł-u

Example (1) illustrates a typical case of paradigmatic derivation: the stems of the motivating adjective and of the derived noun are identical; the difference between the adjective and the noun bears on the distinct sets of inflectional morphemes. By implication, the two words belong to two different paradigms. Typically, the change of paradigm results in a change of word-class. But there are also cases without a change of word-class, such as Polish *fizyk-a* (fem. noun) ‘physics’ vs. *fizyk* (masc. noun) ‘physicist.’ Numerous similar examples can be found in other Slavonic languages, too.

This brings us to the question of how deep are the differences between inflectional and derivational paradigms. No doubt, this question is closely related to that of the differences between inflection and derivation. In contrast to the traditional false belief of a clear-cut boundary between them, cross-linguistic research in recent decades has shown that these differences are not of a disjunctive nature. On the contrary, the boundaries between inflection and derivation are rather fuzzy. The inflection–derivation relations can thus be represented as a cline with more typical (prototypical) and less typical (peripheral) cases along the inflection–derivation continuum (for conclusive evidence cf., for example, Scalise 1988, Dressler 1989, ten Hacken 1994, Booij 2006, Štekauer forthcoming, ten Hacken, this volume).

This situation is necessarily reflected in the relation between inflectional and derivational paradigms. This is the topic of the next section.

20.4 DIFFERENCES AND SIMILARITIES BETWEEN THE INFLECTIONAL AND THE DERIVATIONAL PARADIGMS

One of the most frequently cited differences often held responsible for the “non-existence” of derivational paradigms is the regularity of inflectional paradigms and its absence in word formation. Thus, for example, if any new Slovak noun is classed with the inflectional paradigm *dub* ‘oak’ it acquires all the morphosyntactic features (case, number, gender) and their cumulatively realized formal representations. It is the alleged non-existence of such a regular and automatic assignment of derivational markers to each potential paradigmatic slot that made former discussions of derivational paradigms problematic. For obvious reasons, this is due to, *inter alia*, the absence of competition in the field of inflection, so characteristic for derivational morphology. In particular, there are usually several word-formation strategies available for coining a new complex word. The selection of one of them is determined by both linguistic factors—productivity of word-formation rules, and extralinguistic factors, such as individual preferences affected by one’s age, gender, education, profession, experiences with the language, fashion, etc. (cf. the principle of creativity within productivity constraints in Štekauer 2005; see also Clark, this volume).

In this connection, Pounder (2000: 86) draws attention to the paradigmatic cohesion that is inherently stronger in inflectional paradigms. This, in her view, can be illustrated with suppletion: suppletion exists in inflectional paradigms because of the obligatory nature of inflection. Even if an obligatory path is missing it is supplied by a non-related stem. Word formation is, however, not obligatory and therefore “there is no expectation that a path must be selected to produce a formation for each possible or appropriate semantic rule” (2000: 86) and, for the same reason, there cannot be any defective lexical paradigm: “there is no system-defined number of members. . . it must contain”

(2000: 87). However, contrary to this claim, cases of suppletion in word formation are not rare, as illustrated later in this section.

The same reasoning concerns paradigmatic gaps (or, defective paradigms). Pounder stresses that defectivity is the feature of the lexical (individual) inflectional paradigm if a lexically necessary element is missing. Bauer (1997b: 245) speaks in this connection about the difference in the degree of pattern generalization. He shows that while a partial derivational paradigm such as *manufacture*, *manufacturer*, *manufacturable*, *manufacturability* can be found with other verbal bases too, for example, *break*, *contain*, *drive*, etc., there are gaps like *express* vs. **expresser* (which, obviously, can potentially be filled up).

One of the major issues of research into the paradigmatic nature of morphology concerns its form-oriented or meaning-oriented perspective. The former approach has clearly dominated in the relevant literature, the latter has been either totally ignored or rejected as in the case of Stump (1991). However, the problem with the formal approach is that the paradigm is never a purely formal phenomenon. While it is true that each inflectional paradigm is based on a system of forms, these forms (or, better, the form–meaning units—affixes) express certain cognitively founded categories (case, number, gender, tense, person), etc., that is, subcategories specifying the categories of word classes. Analogically, if we take word-classes as a point of departure for the discussion of derivational paradigms, the relevant subcategories include Agent, Instrument, Action noun, Abstract noun, etc. Stump (1991) argues against such a semantic foundation of derivational paradigms by claiming that while inflection is semantically regular (2):

- (2) sg. *hospital* *fossil* *satire*
 pl. *hospitals* *fossils* *satires*

the semantic relation between the base and the *-ize* derivatives in (3) varies considerably:

- (3) base: *hospital* *fossil* *satire*
 -ize derivative *hospitalize* *fossilize* *satirize*

Stump accounts for the semantic regularity of inflected forms by means of their listedness: while the roots of existing lexemes are listed in the Lexicon, the individual paradigm members are not. By implication, if a paradigm member is not listed its meaning is necessarily predictable by rule. Consequently, it is semantically regular. Stump's argumentation suffers, however, from several deficiencies.

First of all, Stump's example does not appear to be fair to derivation and, by implication, to derivational paradigms. It resembles Chomsky's (1970) demonstration of the regular and productive syntax vs. the idiosyncratic and unproductive derivation which was called into question by, for example, Di Sciullo and Williams (1987) and Štekauer (1998). Crucially, the inflectional paradigm is based on the semantic relationship, for example, the cognitive category of PLURALITY, no matter what the formal realization of this category is in the individual cases (cf. a number of different plural morphemes

in Slovak: *-i, -y, -ovia, -e, -á, -ia, -tá*). Therefore, “fair” argumentation also requires us to put the relation between the members of a minimal (i.e. two-member) derivational paradigm on a truly semantic basis by saying that the paradigm rests on the cognitive category of, for example, RESULT OF ACTION (which may, in English, be formally represented by several prefixes and suffixes). Analogically to a number of different inflectional paradigms in inflectional languages (Slovak, for instance, has twelve different paradigms for nouns) one can postulate different paradigms for the individual derivational categories, such as the RESULT OF ACTION, AGENT, INSTRUMENT, etc. depending on their formal realization. By implication, one can speak about several distinct derivational paradigms within the category of AGENT, defined by the relation of a verbal base and the suffix *-er*; nominal base and the suffix *-ist*; nominal base and the suffix *-ian*; nominal base and the semisuffix *-man*; verbal base and the suffix *-ee*; etc. Each of these paradigms functions both as a static pattern and a dynamic tool for the derivation of analogical coinages. In this way, the principle of regularity and predictability is also introduced into the system of derivational paradigms. (The role of actual and potential derivatives will be discussed later.) This view receives important support from the phenomenon of suppletion. Suppletion exists both in inflection and derivation as illustrated with the paradigmatic pairs like *sound—sonic*, *ear—auditive*, *sun—solar*, and a number of others. The individual members of these paradigms are not held together formally but semantically—in the same way as the members of the corresponding inflectional paradigms. In fact, this is a natural consequence of Stump’s conception of paradigm function as “a function from the root of a lexeme to one of the fully inflected words in the paradigm of that lexeme.” Translated into my proposal, derivational paradigmatic relations are based on formal realization of a cognitive category by an affixation process. This cognitively founded process relates the stem as an input and the resulting derived word as an output.

Another instance of analogy between inflectional and derivational paradigms is obtained from a series of paradigmatic function relations. They are characteristic for both inflectional and derivational paradigms. Stump provides an example (1991: 683) from Breton of the transition from the root *lenn* ‘read’ to the word *lenn-f-e-mp* ‘we would read’ by application of three successive inflectional rules. The paradigm function thus comprises a sequence of inflectional rules. There are numerous analogical instances in derivation where the paradigm function relating the stem and the relevant complex word comprises a sequence of derivational rules:

- (4) a. *institute-ion-al-ize-ation*
 b. *list* → *lístek* → *lístkový* → *lístkovitý* → *lístkovitost*
 ‘leaf’ ‘leaflet_N’ ‘leaflet_A’ ‘characterized by leaflets_A’ ‘property of having leaflets_N’
 (Dokulil 1962)

It follows from what we have said that there are no clearcut boundaries between inflectional paradigms and derivational paradigms, in the same way as there is no clearcut boundary between inflection and derivation in general.

More evidence of the similar nature of inflectional and derivational paradigms comes from the meaning modifications that occur both in derivation and inflection. Therefore, parallel to the meaning differences within the above mentioned category of RESULT OF ACTION (Location—*hospitalize*; Manner—*satirize*; Effect—*fossilize*) there are instances of meaning modifications in the exemplified inflectional paradigm of PLURALITY (cf. duality, collectivity, mass nouns, etc.) in a cross-linguistic perspective.

Furthermore, the argument in favor of the semantic regularity of inflected forms (i.e. non-listed and hence not undergoing any meaning modifications) are also applicable to highly productive derivational rules. Thus, there are examples of absolutely productive and therefore fully predictable paradigms in derivation. The frequently quoted examples include the derivation of verbal nouns in *-ing* in English, and the full predictability of English nominalizations derived from *-ize* (*-ation*), *-ify* (*-ification*), and *-ate* (*-ion*) verbs.

From the opposite end, there are numerous and well-known exceptions from the absolute regularity of inflectional paradigms, such as incomplete paradigms (singularia tantum, pluralia tantum, collectives, mass nouns, etc.), irregularly and therefore unpredictably marked members of paradigms (e.g. absence of third person singular of modal verbs, foreign plural of nouns, cases of suppletion, absence of formal expression of certain morphosyntactic categories due to semantic reasons).

The previous discussion has demonstrated that there are evidently no sharp boundaries between the two types of paradigm, and therefore we should seek for the features they share. Bauer (1997b: 245ff.) when checking the derivational relevance of van Marle's (1994) inflectional paradigm criteria concludes that what is common to both types of paradigm is that they are organized round the basic form, that is, they share a common base irrespective of its specific form. Bauer (1997b: 247) shows that in some cases the base must be reconstructed. For example, for *aggressor* where there is no base *aggress*, it must be reconstructed from the paradigm, including *aggressor*, *aggressive*, *aggression*, and their derivatives. He points out that the ability to reconstruct a base from a paradigm is a feature identical to inflectional paradigms.

Moreover, the paradigm pressure, so typical of inflection, can also be seen in the formation of neologisms which due to the non-automatic nature of derivation are not perceived as such. Bauer points out that "[i]f we invent a new word like *Kuwaitize*. . . then *Kuwaitization* is no more (or less) striking than the verb from which it is superficially derived" (1997b: 250). Similarly, Horecký et al. (1989: 42) point out that new complex words do not come into existence as isolated units but rather as complete paradigms. So, for example, if a new field of science is established, the coining of its name motivates the coming-into-existence of other vitally needed derivational "cases", including the name for an expert in the field, one or more adjectives, possibly an adverb, as illustrated by Horecký et al. (1989) in (5):

- (5) *prognóza* 'prognosis' >
 prognost-ika 'science of forecasting'
 *prognost-ický*_{Adj} 'related to the science of forecasting'
 *prognosticky*_{Adv} 'by using the science of forecasting'

prognost-ik ‘forecaster’
prognóz-ovat ‘to forecast’
*prognózova-nie*_N ‘forecasting’
*prognost-ický*_{Adj} ‘related to forecasting’
prognóz-ny ‘related to forecast’—*prognózn-e* ‘by way of forecast’

By implication, an important feature of the derivational paradigm is its capacity to show the as yet unactualized potential, the current gaps which may be, in the course of language development, filled with actual words.

Consequently, the neologism is experienced as an actualization of existing possibilities, just as is the case in inflectional paradigms. This seems to be a strong point in comparing inflectional and derivational paradigms. Both of them rest on a set of slots representing certain categories that can potentially be actualized. While the degree of actualization of potentiality in inflectional morphology is very high, the percentage of the actualized potential slots in derivational morphology is much lower. But, given the vague boundaries between the actual and the potential words, one finds it advantageous (or even necessary) to extend the scope of derivational paradigms to encompass potential words too. These reasons fully justify Bauer’s assumption (1997b: 253) that the availability of slots is more important for the paradigm than the forms which fill them. This view appears as early as Zemskaja (1978) who proposes to use the term “type of derivational paradigm” which abstracts from concrete realizations of the individual slots.

The significance of this position is that forms like suppletivism are not an obstacle to the existence of derivational paradigms and that pairs like *sound—sonic*, *ear—auditive* are in a paradigmatic relation. This is, in fact, also assumed by Stump (1991) when he says that

paradigm function is assumed to link the root *x* of a lexeme *L* with a word *w* in *L*’s paradigm even in those cases in which *w* doesn’t come from *x* through the operation of any morpholexical rule (inflectional or otherwise). One such case is when *x* and *w* stand in a suppletive relation. . . . (Stump 1991: 684)

By implication, if formal unity is not crucial, what gains in significance is the unity established by the semantics of cognitive category.

20.5 APPROACHES TO THE DERIVATIONAL PARADIGM

This section presents various approaches to the concept of a derivational paradigm. What should be stressed at the very beginning (and what is sometimes ignored in the literature) is that the notion of a derivational paradigm has two basic meanings (as already

indicated above). First, it may refer to a system of relationships covering both actual and potential words (static aspect). Second, it may function as a device for the formation of new complex words (dynamic aspect).

20.5.1 Individual Cases

Booij's treatment of the derivational paradigm (1997, 2008) concentrates on its dynamic aspect when he says that "in many cases new complex words are created on the basis of relations between existing words" (2008: 29). A case in point are toponymic derivations in Dutch (Booij 1997: 45ff.) showing regular paradigmatic relations between the toponym, the name of the inhabitant, the derived adjective, and the name of a female inhabitant:

(6) Toponym	Inhabitant	Adjective	Female inhabitant
<i>België</i>	<i>Belg</i>	<i>Belg-isch</i>	<i>Belg-isch-e</i>
<i>Finland</i>	<i>Fin</i>	<i>Fin-s</i>	<i>Fin-s-e</i>
<i>Hongarije</i>	<i>Hongaar</i>	<i>Hongaar-s</i>	<i>Hongaar-s-e</i>
<i>Noorwegen</i>	<i>Noor</i>	<i>Noor-s</i>	<i>Noor-s-e</i>
<i>Rusland</i>	<i>Rus</i>	<i>Russ-isch</i>	<i>Russ-isch-e</i>

Booij demonstrates the interrelation between these names. In particular, he shows that the adjectival suffixes *-s* and *-isch* are attached to a base which is formally identical to the name of the inhabitant. Female names are not derived from the neutral counterparts. Instead, they are derived from the corresponding toponymic adjectives by attaching the suffix *-e*. It follows from the example that there is no single form which can serve as a common base for the derivation of the other names.

Booij's arguments in support of the relevance of the concept of derivational paradigm can be summarized as follows (Booij 2008: 29ff):

- Idiosyncratic properties of existing complex words recur in a larger complex word, e.g. the idiosyncratic meaning of the deverbal Dutch noun *woning* derived from the verb *woon* 'to live' by means of the suffix *-ing* used to derive deverbal action nouns, recurs in all words derived from *woning*, e.g. *woningnood* 'shortage of houses'.
- Paradigmatic word formation, such as the formation of female nouns in *-ster* as discussed in van Marle (1985).
- Bracketing paradoxes can be eliminated by postulating the paradigmatic relationship as in the case of *transformational grammarian*:

$$\text{grammar} : \text{grammarian} = \text{transformational grammar} : X$$
 where *X* is *transformational grammarian*
- Cases of hypercharacterization where a word ends in a suffix that does not contribute an additional meaning, e.g. Dutch *Dominicaner* 'Dominican' is derived

from *Dominicaan* which has the same meaning. The suffix *-er* has been added by analogy to the general rule of personal noun formation.

- Relations between words in *-ist* 'adherent of ideology' and *-ism* 'ideology'. If they are accounted by paradigmatic relationship, Booij claims, it is not necessary to explain them by the process of truncation. Furthermore, we do not have to decide on the direction of derivation.
- Finally, he refers to Krott (2001) who shows that the use of the linking element in Dutch compounds depends on paradigmatic relations, i.e. compounds that begin or end in the same constituent.

20.5.2 System-based Treatment

20.5.2.1 *Dokulil, Horecký, and Their Tradition*

The discussion of derivational paradigms has been inherent in the onomasiological tradition of word formation (e.g. Dokulil 1962, Buzássyová 1974, Horecký et al. 1989, Furdík 2004). Derivational paradigms in a narrow sense are viewed as just one level of highly complex derivational relationships within the whole word-formation system. All the paradigmatic relationships are examined on the basis of the derivational principle of motivation. It is postulated that the derivational system is based on paradigmatic relations reflecting the internal structure of the word-formation system of a language as a whole and that word-formation relations cannot be reduced to those between the motivating and the motivated words. Rather, each such pair of words is the basis for much more complex relations. One of them is the derivational series whose neighboring members are successively related by motivation:

- (7) *písať* 'to write' → *za-písať* 'to record by writing' → *zapis-ovať* 'to keep recording by writing' → *zapisova-tel'* 'registrar/recorder' → *zapisovatel'-ka* 'female registrar/recorder' → *zapisovatel'k-in* 'belonging to female registrar, recorder' (Furdík 2004: 73)

The derivational paradigm in the narrow sense is conceived as an ordered set of all complex words directly derived from a single basic (motivating) word (Horecký et al. 1989: 38, Furdík 2004: 73ff.). Importantly, the motivated words do not mutually motivate one another. The idea is illustrated in (8):

(8) <i>škola</i> 'school'	<i>škol-ák</i>	'schoolboy'
	<i>škol-ník</i>	'school janitor'
	<i>škôl-ka</i>	'kindergarten'
	<i>škol-stvo</i>	'education system'
	<i>škol-ička</i>	'small school' (Furdík 2004: 74)

As pointed out by Furdík (2004) these kinds of sets of co-motivated words form a sort of system of “derivational cases” even if these cases are not as systematic as inflectional cases.

Dokulil (1962: 14) lays emphasis on the relation between the static and the dynamic aspects of the derivational paradigm: he conceives of it as a pattern established by a sequence of derivational operations which makes it possible for other words to form their derivatives:

(9) <i>list</i> ‘leaf’	→	<i>líst-ek</i>	→	<i>lístk-ový</i>	→	<i>lístkov-itý</i> ‘characterized by ‘leaflet’ _N ‘leaflet’ _A leaflets’ _A
<i>květ</i> ‘flower’	→	<i>kvít-ek</i>	→	<i>kvítk-ový</i>	→	<i>kvítkov-itý</i>
<i>zub</i> ‘tooth’	→	<i>zoub-ek</i>	→	<i>zoubk-ový</i>	→	<i>zoubkov-itý</i>
<i>roh</i> ‘corner’	→	<i>růž-ek</i>	→	<i>růžk-ový</i>	→	<i>růžkov-itý</i>
<i>smrk</i> ‘spruce’	→	<i>smrč-ek</i>	→	<i>smrčk-ový</i>	→	<i>smrčkov-itý</i> etc.

Finally, combinations of word-formation series and word-formation paradigms establish the derivational nest (or cluster) constituted, syntagmatically, by a set of derivational series (chains) and, paradigmatically, by an ordered set of words with the same root interrelated by synchronic relations of derivation (Horecký et al. 1989: 38).

The relationships between the individual words in the derivational nest are based on a common root morpheme and a common onomasiological mark. The capacity of the individual members of the derivational nest to motivate other words, and thus contribute to the development of the paradigmatic relationships, depends on a range of factors, including word-class, degree of their own motivation, degree of semantic correlation with its motivating word, the complexity of its semantic structure, its derivational valence, and productivity of the corresponding word-formation type.

20.5.2.2 *Van Marle*

Van Marle (1985), too, discusses derivational paradigms within the broader context of paradigmatic foundations of word-formation. Against the background of Saussurean structuralism, van Marle (1985: 83ff.) emphasizes the twofold nature of the study of paradigmatics: the study of various types of entities and the study of various relationships between these entities. One such paradigmatic relationship is the relationship between different types of complex words (rather than between the base and its derivative). He illustrates this with the relationship between relativizing complex adjectives in Dutch—one type derived by *-erig* (stronger relativization), the other by *-ig*: *groen-erig* vs. *groen-ig* (both of them meaning ‘greenish’). The importance of paradigmatic relationships between sets of complex words like these consists, as emphasized by van Marle, in the fact that “the semantics of these sets of words cannot be fully understood without taking the semantics of the other into consideration” (1985: 86).

One of the basic terms used in van Marle’s theory is *morphological category* understood as “a series of words sharing identical formal feature and identical semantic feature” (1985: 88). As an example, all Dutch deadjectival adjectives in *-ig* (*groen-ig*

‘greenish’, *kal-ig* ‘baldish’, *natt-ig* ‘wettish’) having the meaning of “subjective relativization” represent a morphological category. The common feature of meaning bound to the common formal feature of a series of words is called *categorial value*. Then, by *morphological structure* van Marle understands “the presence of a categorial value which corresponds to a certain constant element of form” (1985: 128) (e.g. Dutch words in *-ig* whose meaning is “relativization” of the adjective).

The relationships between various morphological categories are labeled as *intercategorial relationships*. Van Marle distinguishes four types of *intercategorial relationships* between two or more morphological categories:

- categories with an identical base on the word-class level;
- categories displaying related categorial value (i.e. they are characterized by semantic similarities) (cf. the suffixes *-ig* and *-erig* in Dutch showing semantic similarity; they are used to form “subjectively relativizing” adjectives, and differ just in the degree of relativization);
- categories displaying identical categorial value (i.e. they are characterized by semantic identity)—these are rival morphological categories;
- categories sharing an identical affix—one and the same suffix may figure in different morphological categories

The sum total of the relationships of a given morphological category constitutes the *derivational domain* of this category. While morphological category is predominantly an entitative concept identifying “building blocks” of morphological systems, *word-type* is its dynamic counterpart which underlies the formation of new complex words. Finally, *morphological systems* are defined as systems composed of word-types.

It should be noted that the relationship of semantic identity of two or more morphological categories is, in the above example, absent, because the *-ig* category does not have any rival category.

Moreover, the example shows that the relationship of semantic similarity is gradual. The categorial value of both *-ig* and *-erig* categories is “subjectively relativizing” (of different degree) while the categorial value of the *-achtig* category is “objectively relativizing.”

All in all, van Marle’s approach emphasizes paradigmatic foundations of the whole system of word-formation in terms of both word-formation-analysis and word-formation proper.

20.5.2.3 *Stump*

Stump (1991) derives his theory of derivational paradigms from that of inflectional paradigms. His understanding of the inflectional paradigm follows from his *hypothesis of paradigm-based inflectional semantics* (1991: 679). The basic units of his theory include the word as a fully inflected expression which constitutes an inflectional paradigm; a set of morphosyntactic features σ ; morpholexical rules (MLR); and the *paradigm function* defined as “a function from the root of a lexeme to one of the fully inflected words

in the paradigm of that lexeme” (1991: 683). An important difference between paradigm functions and inflectional morpholexical rules can be found in languages with rich morphology where the transition from root to word may include several steps, that is, the successive application of two or more rules. Stump’s example from Breton is illustrative. The transition from the root *lenn* ‘read’ to the word *lenn-f-e-mp* ‘we would read’ includes the application of three inflectional rules. Then, “the paradigm function that links the root to the fully inflected word in such cases. . . comprises a sequence of such rules” (1991: 683).

Stump applies his theory of paradigm functions to derivational paradigms which are based on a binary relation: “Every derivational rule *R* defines a simple two-member paradigm consisting of an input expression (which *R* takes as its argument) and the corresponding output expression (the value of *R* for that argument)” (1991: 708).

Stump (1991: 721ff) argues that the relationship between the two members of a derivational paradigm is a formal one, which means that “the paradigm function F_{ist} linking *art* to *artist* is distinct from the function f_{er} linking *jewel* to *jeweler*.” He rejects a semantically based derivational paradigm such as $f_{personal:noun}$ relating both *art* to *artist* and *jewel* to *jeweler*. Such a derivational paradigm function would be realized by a range of different morpholexical rules, including MLR_{ist} , MLR_{er} , etc.

The semantic approach should, in his view, be rejected because it would make derivational paradigm functions (DPF) starkly different from inflectional paradigm functions. One of the reasons is the existence of derivational doublets like *basal/basic*, *conformance/conformity*, *conformer/conformist*, *elemental/elementary*, *to legitimate/legitimize*, *variance/variation*, etc. He maintains that a proponent of a semantic approach has two options with these doublets. First, derivational paradigm functions would not be functions but relations so that a personal noun would relate [_V*inform*] to both [_N*informer*] and [_N*informant*]. Secondly, each such doublet signals the existence of two contrasting DPFs with distinct semantic definitions. This view would entail a large number of distinct DPFs, one for relating *conform* to *conformer*, one for *conform* to *conformist*, a different one for the pair *breathe/breather*, another one for *coast/coaster*, still a different one for *twist/twister*. What he claims is that this approach “would therefore entail a stock of DPFs whose vastness and idiosyncrasy would rival that of the lexicon” (Stump 1991: 722).

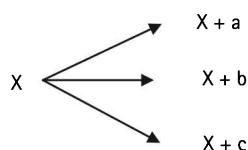
Stump believes that his formal conception provides a uniform theory of paradigm functions. His Derivational Paradigm Functions provide the same sort of basis for the assignment of logical representations to derived forms as Inflectional Paradigm Functions (IPFs) do in the case of inflected forms: the Elsewhere Condition applies indiscriminately in the evaluation of both IPFs and DPFs; the members of pairs such as *conform/conformer*, *breathe/breather*, *coast/coaster*, and *twist/twister* are linked by a single DPF f_{er} , and their semantic differences are taken to reflect stipulations in their lexical entries rather than definitional differences among contrasting DPFs.

20.5.2.4 Pounder

Pounder’s model is inspired by Guilbert (1975). Guilbert (like Bauer and Horecký et al.) includes in the paradigm potential formations. He does not take semantics into

consideration, and therefore the formal representation of paradigm is exclusively based on a combination of stem and affix, as illustrated in (10) where X is the base and a, b, c the affixes:

(10)



Furthermore, Guilbert points out that (a) the stem may be morphologically complex; (b) each meaning of a polysemous stem underlies a different paradigm; (c) there are two basic types of derivational paradigms:

(a) Fan-like as illustrated in (10)

(b) Cumulative type as illustrated in (11):

(11) X → X + a → (X + a) + b
 play → playful → playfulness

Pounder proposes her own model of the derivational paradigm, the so-called *process-and-paradigm* model. She views the derivational paradigm as a *dynamic* model for the formation of new lexemes based on word-formation rules which fill in the potential slots in a morphological paradigm to form a complex lexeme. Analogically to her distinction between the lexical (i.e. individual) inflectional paradigm and the morphological inflectional paradigm, she draws a distinction between the *systemic morphological paradigm* conceived of as a pattern for a particular class of complex words, or, more specifically, “as a set of possible paths defined by series of operations applicable to a given base of a given lexico-syntactic category producing correct/possible complex lexemes” (2000: 95), and the *lexical paradigm* as a specific application of this pattern covering “the set of word-formations with a common lexeme base,” or, as “a structured pattern of ‘instructions’ for operations on stems, where feasible” (2000: 85), with the emphasis being laid on the relations within the paradigm. The lexicon only contains those paths of the systemic word-formation paradigms that are actually used. Given the fact that one of the characteristic features of the paradigm is the potential for synonymy and polysemy relations and that derivational morphology is based on competition (unlike inflectional morphology), individual operations in the systemic paradigm have unequal significance, depending on their productivity. As a result, “only a subset of the morphological paradigm is operative: from each synonymy set, generally, only a subset is actively used, and some synonymy sets are not used at all. . .” (2000: 641).

As an intermediate level between these two types of derivational paradigm, Pounder proposes the so-called *class-specific paradigm* as a subparadigm of the systemic

paradigm which is productive in the particular lexicosemantic class. Consequently, it differs from the systemic paradigm only in size. From this point of view, the lexical paradigm may be viewed as an instantiation of the class-specific paradigm in the lexicon for a given base.

Like many other approaches to the derivational paradigm, Pounder takes the *common stem* as a point of departure. An important difference from Guilbert is that Pounder assumes that the derivational paradigm should not be restricted to affixation processes. Furthermore, apart from the common stem, each paradigm can be defined in terms of four criteria: (a) word-class (all verb formations together, all adjective formations together, etc.); (b) common semantic rule; (c) secondary formal rules (if relevant)—rather than constituting a new word-formation meaning they just modify the existing one, such as the functions connected with evaluative morphology; (d) lexical paradigm, including information, such as regular elements of meaning (“habitual,” “human,” “object”, etc.; syntactic features of the complex word; stylistic marking, and expressivity).

Pounder’s approach gives support to the significance of potential words in the system of derivational paradigms: it is not possible to arrive at a full understanding of a word-formation system in diachrony or synchrony unless one takes into account the full set of derivational paths a base may follow (2000: 662).

20.6 CONCLUSIONS

The overview of various approaches to the concept of derivational paradigm makes it possible to draw the following conclusions concerning their defining characteristics:

- The derivational paradigm is a well-justified concept reflecting one dimension in the multidimensional network of relationships in the complex system of word formation.
- The derivational paradigm manifests itself on two axes. First, horizontally, it is constituted within a particular cognitive category as a binary relation sharing a common stem. A single cognitive category may include several paradigms (the derivational paradigm of AGENTS includes several different paradigms that can be exemplified with *teach—teacher*, *journal—journalist*, *library—librarian*, *escape—escapee*, *milk—milkman*, etc.). This binary type of paradigm is extendable as a series of successively potentiated operations unified by a common root (base), governed by relevant cognitive categories (examples (6) and (9) above). In both of these cases, a derivational paradigm represents a pattern for analogical formations falling within the same paradigm (examples (8) and (11) above). Second, vertically, the derivational paradigm represents a range of relations unified by a common stem (examples (7) and (10) above).
- Inflectional and derivational paradigms share many characteristics; the similarity between these two types of paradigm seems to be closer than formerly believed.

- Potentiality is a crucial term for the concept of derivational paradigm. It guarantees a high level of predictability and regularity (approaching that of the inflectional paradigm) in the sense that the existing gaps in the system can be filled anytime with actual words.
- It is for this reason that one of the essential features of the derivational paradigm is the availability of slots (filled with potential words) that are more important for the paradigm than the forms which fill them.
- The derivational paradigm has two interrelated facets—it represents both the static system of relations and the dynamic system actively contributing to the extension of the lexicon by functioning as a pattern for word-formation.
- It may be assumed that the concept of the derivational paradigm is restricted to affixation in terms of word-formation processes participating in its constitution. This follows from the requirement of systematic, regular, and predictable relationships within the paradigm. Conceived this way, the derivational paradigm encompasses all the potential affixations of a particular common root/stem. This system is a closed one in terms of the number of potential complex words. In this respect it is close to the concept of the inflectional paradigm. Conceived broadly, the idea of a derivational paradigm, constituted by all word-formation processes, is paradigmatically vacuous because it does not lead to a predictable and regularly organized system of complex words. A system of complex words produced by all the word-formation processes and related by a common stem is an open system, different from the closed system of paradigm.

CHAPTER 21

AFFIX ORDERING IN DERIVATION

PAULIINA SAARINEN AND JENNIFER HAY

21.1 INTRODUCTION

THIS chapter concerns factors influencing the order in which derivational affixes may attach to a word. Why, for example, is “*reddishness*” so much better than “*rednessish*”? There are a variety of factors which appear to influence affix-ordering—each of which can be observed across multiple languages. However, as has been comprehensively argued by Manova and Aronoff (2010), while there are observable cross-linguistic tendencies, there is certainly no existing recipe of factors that can be shown to dictate all languages’ affix-ordering behaviors. Nor should we really expect this. The object of study is inherently different across different languages due to the frequency and productivity distributions of the affixes, the semantic notions expressed by them, the factors that lend themselves to parsing affixes in speech perception, the statistical (ir)regularities, and the degree of semantic and phonological transparency. Different languages balance these different tensions differently, and have their own solutions to the problem of establishing an “optimal” affix order.

Thus, the task of outlining affix-ordering behavior is somewhat complex, as each language presents its own inventory of affixes, and its own unique package of factors and influences which together constrain its affix ordering behavior. While there are common threads and tendencies, reports on individual languages vary according to factors such as the apparent relative importance of semantics and phonology, the degree to which psycholinguistic factors are at play, whether affixes appear to fall into discrete categories, the transitivity of ordering relations, and possibilities for recursion. In his work on Bantu languages, Hyman (2003: 246) argues that affix systems vary, and that “each suffix system represents a language-specific resolution of a basic tension between two competing pressures: the pressure for affix ordering to be compositional vs. the pressure for affix ordering to be fixed (invariant).”

Given the considerable language-specificity in how the various factors are dealt with, this short overview is necessarily far from exhaustive. Our overview, like the literature, contains a considerable overrepresentation of data and phenomena from English. The problem of affix-ordering in English has been the object of an impressive amount of discussion and theorizing, which is certainly disproportionate in relation to work on affix-ordering in other languages.

21.2 INFLECTION AND DERIVATION

Although this overview focuses on derivational affixation, we start by saying a few words about the ordering of derivational affixes with respect to other forms of word formation. In general, derivational affixation occurs closer to the stem than other forms of word formation, such as inflectional affixation and compounding (e.g. Greenberg 1963b, Anderson 1982). It is very unusual cross-linguistically for derivational affixes to attach to inflected forms. There are exceptions to this tendency. In Athapaskan languages, for example, the inflectional affixes appear in an order which is interspersed with derivational affixes (Rice 2000). An explanation often invoked for the (usually) systematic ordering of inflectional and derivational affixation is that affixes attach in an incremental manner, and in an order which reflects syntactic function. Viewed this way, the inflectional affixes—being more “syntactic” than derivational affixes, are usually the last to be added.

Baker (1985a: 375) directly ties affix ordering to syntactic order, arguing that “[m]orphological derivations must directly reflect syntactic derivations (and vice versa).” This correspondence between syntax and morphology is known as the Mirror Principle and has been argued to be “... an exceptionless generalization, with strong empirical content given the constraints on word formation” (Alsina 1999: 6). This claim is too strong, however. There does seem to be a general correspondence between syntactic and morphological order in languages (and especially amongst inflectional affixes, which play a more syntactic role in general) but this is far from universal (see, e.g., examples and discussion in Hyman 2003).

21.3 AFFIX CO-OCCURRENCE RESTRICTIONS

When considering what factors constrain the *overall* ordering of affixes in a language, much of the detailed answer comes down to description of more local co-occurrence patterns regarding particular pairs of affixes, or relationships between certain types of affixes and certain types of bases.

The term “potentiation” refers to particular affix combinations that co-occur with high frequency, where the addition of the first affix makes possible/likely the

addition of the second. Thus an affix does not only place restrictions on what can attach to it, but it can also make particular other affix additions particularly likely. An example of potentiation in English is the relationship between the prefix *en-* and the suffix *-ment*. *En-* potentiates *-ment*, in that the addition of this particular prefix to a stem makes the suffix also eligible to attach (cf. *enchantment*, *enjoyment*, *enlightenment*, etc.)

There are also many restrictions in the other direction—restrictions about which affixes can occur together. For example, some suffixes have been claimed to only occur after one particular other suffix (e.g. *-ic* in *modernistic* or denominal *-er* in *vacationer*), and many do not allow any suffixes to precede them—they can only attach to unsuffixed base words (e.g. *-hood* as in *adulthood* or *-ish* as in *reddish*) (Fabb 1988). Such restrictions originally tended to be phrased in terms of restrictions driven by the affix, about what it could attach to—so-called *affix-driven* selectional restrictions. Many of these have subsequently been argued to be *base-driven* selectional restrictions—restrictions that the base form places on what can attach to it (see, e.g., Giegerich 1999, Plag 1999). That is, any co-occurrence restrictions can be framed in terms of restrictions that an already-affixed *base* places on what can be added to it. Framed differently, affixes have restrictions about what can attach to them, rather than what they can attach to. These restrictions are most often phonological or semantic. Aronoff and Fuhrhop (2002), for example, designate a class of suffixes to be “Closing Suffixes”—they do not allow any further derivational suffixes to be added. In German, a derivational closing suffix *-ling*, for example, cannot take another derivational suffix (although the attachment of one or more inflectional suffixes is possible).

Whatever the origin and type of the selectional restrictions, they tend to be exceedingly local; restrictions operate on the co-occurrence of two affixes joining in adjacent cycles. This property makes the morphological system well-modeled as a series of weighted bigram constraints, basically a bank of pairwise co-occurrence restrictions (Ryan 2010).

21.4 SEMANTIC FACTORS

The overall order of affixes within a word is often constrained by semantic factors. Bybee (1985) surveyed over fifty languages, looking at the order with which affixes attach to a verb. She argues that there is a general “Relevance Hierarchy” which states that the more semantically relevant a morpheme is to the verb, the closer to the verb it is placed. This generalization captures the ordering of derivation inside inflection, as derivational affixes have a much more radical effect on the semantics of the stem. It also places constraints on the relative order of affixes within this class. For example, her survey reveals the following hierarchy in terms of the order of verbal markers: *Valence* > *Voice* > *Aspect* > *Tense* > *Mood* > *Agreement*. She claims this hierarchy reflects the degree of semantic relevance to the verb.

The semantics of words can constrain affix ordering because of the apparent scopal relationships invoked by different orderings. Each suffix that attaches to the base can add a layer of semantic meaning to the preceding elements, as if the words are built incrementally step by step. This is also known as layered or hierarchical morphology. For instance, if suffixes A and B are ordered BASE+AB, the suffix B has scope over BASE+A because the meaning of B alters the meaning of BASE+A. In some languages, the reverse order BASE+BA can also occur. In Yup'ik, for example, variable affix order is possible in some words to express different scopal relationships (Mithun 1999). Example (1) shows how the semantics of the whole expression AB or BA is derived from the semantics of the suffix A and the suffix B compositionally. In the complex noun on the left, the outermost suffix *-cuar* 'little' is the modifier to the derived stem *yug-pag* 'giant.' In the noun on the right, the outermost suffix *-pag* 'big' modifies the derived stem *yug-cuar* 'midget.'

(1)

yug-pag-cuar	yug-cuar-pag
person-big-little	person-little-big
'little giant'	'big midget'

(Mithun 1999: 43)

Also in Kwakw'ala, a Wakashan language spoken in northern Vancouver Island and adjacent mainland British Columbia, there is a relative freedom in the ordering of suffixes. Since in Kwakw'ala only semantics (scope) in general poses limits to suffix co-occurrence and suffixes themselves are not sensitive to the complexity of the bases they attach to (Anderson 1985: 25; also Štekauer et al. 2012: 147), distinct meanings can be created by stacking suffixes in specific orders. As shown in (2), the relative ordering of the suffixes *-exsd* 'want' and *-amas* 'cause' creates differences in scope:

(2) ne'nakw	'go home'
ne'nakw'exsd	'want to go home'
ne'nakw'exsdamas	'cause to want to go home'

The most comprehensive and detailed study of semantic scope is that of Rice (2000) on Athapaskan languages. Rice argues that morpheme order in Athapaskan verbs is constrained by semantic scope, which is expressed by or reflected in linear order. Semantic scope manifests itself both in fixed or variable orders. When the scopal affix order is fixed, the elements occur in a fixed order in relation to each other. When the scopal affix order is variable, the order of elements can reverse and the interpretation changes based on the order. As predicted by Rice's "semantic scope" hypothesis, either order is attested when no semantic relationship can be established between the affixes.

Fixed orderings are not always semantically driven. In the Yenissean language Ket, for example, affixes fall into classes based on their grammatical category and they are

positioned in a predetermined order based on a rigid template. The template is arbitrary with no apparent semantic relationship between the classes. Ket verbs vary in terms of the number of position classes they can take. Štekauer et al. (2012) identify eight derivational and inflectional position classes in Ket for prefixes alone.

In terms of the types of semantic relationships created by affixation, affixes are sometimes categorized into different groups. These categories differ in terms of semantic “closeness” to the verb and semantic transparency, as well as in many other non-semantic properties. When languages have been posited to have discrete categories of derivational affixes, these categories, or “classes” have assumed central theoretical importance, as discussed in the next section.

21.5 “MULTIPLE STRATA” MODELS

As outlined in the previous section, affixes displaying different degrees of semantic transparency are often hypothesized to belong to different classes, or levels. These different levels are often referred to as “strata.” The term “strata” originally referred simply to the distinction between native phonology and loan word phonology (e.g. Saciuk 1969). In English, for example, the term referred to the difference between Germanic and Romance vocabulary. Affixes in these two groups tend to display divergent behaviors with respect to the phonological and semantic relationships to their bases.

In subsequent work, it is these different behaviors, rather than the etymological origins of affixes, that are the primary drivers for division into two groups. In terms of phonology, affixes have typically been clustered into two groups. One group of affixes can be semantically and phonologically opaque, and can cause a stress-shifting behavior in the base (cf. *insane/insanity*). The other group of affixes tend to retain phonological and semantic transparency, and do not trigger stress-shifting (e.g. *insane/insaneness*).

The precise criteria and motivation for the division into categories, and the nomenclature for the categories have varied across different works. The first general category can be referred to as weak, non-neutral, stress-shifting (e.g. Chomsky and Halle 1968, Siegel 1974), level 1 (e.g. Selkirk 1982, Kiparsky 1982c, Mohanan 1986), or class 1 (Siegel 1974) affixes. The second are referred to as strong, neutral, non-stress-shifting, level 2, or class 2 affixes respectively. The choice of term is somewhat dependent on the point in history, the exact motivation and criteria for division into classes, and the theoretical context of the discussion.

In *The Sound Pattern of English*, Chomsky and Halle (1968) identified two kinds of morpheme boundaries, strong (#) boundaries and weak (+) boundaries. Strong #-affixes (like *-ness*) do not affect the stress of the words to which they attach, and are also known as “neutral” affixes. By contrast, +-affixes (like *-ity*) form a weak (non-neutral) boundary which triggers a stress-pattern change in the base. Non-neutral affixes can also affect other phonological material in the base, whereas neutral affixes do not.

According to Siegel (1974), each suffix creates only one type of boundary. In order to account for the different phonological and—particularly—stress related behaviors, it is posited that the non-neutral affixes attach to the base first. A set of phonological rules then applies, and stress is assigned. After this process, the neutral affixes are added. This idea of different stages in the derivation elegantly accounts for the generalization that the non-neutral affixes are subject to phonological rules, and for the observation that neutral affixes seem to occur outside non-neutral affixes, but not vice-versa (Siegel 1974).

Indeed, the observation that there are different classes of affixes has been quite theoretically attractive, as it can easily be incorporated into theoretical machinery in order to elegantly account for restrictions on the relative order of affix classes. That is, if we simply assume that affixes belonging to different classes attach at different stages in the derivation process, then this would obviously constrain the relative ordering in which they can appear. Selkirk (1982) dubbed this general insight regarding the relative ordering of two classes of affixes the “Affix Ordering Generalisation.” It posits that English prefixes and suffixes belong to either of two levels (or strata) and that an affix can only be attached to an affix of the same or lower level. This is also known as level-ordering and adopted into most models of lexical phonology (e.g. Kiparsky 1982c, Mohanan 1986). Affixes are associated with one of the two strata. For example:

(3)

Class 1 suffixes: +ion, +ity, +y, +al, +ate, +ous, +ive, +able, +ize

Class 1 prefixes: re+, con+, de+, sub+, pre+, in+, en+, be+

Class 2 suffixes: #ness, #less, #hood, #ful, #ly, #y, #like, #ist, #able, #ize

Class 2 prefixes: re#, sub#, un#, non#, de#, semi#, anti#

(from Spencer 1991: 79)

A number of languages have received stratum-based analyses. Booij (2001) identifies Dutch, Malayam, Dakota, Mandarin, Icelandic, German, Igbo, Italian, Basque, and Turkish as languages which have been analyzed as containing multiple levels of derivation. Authors who are particularly associated with the use of stratum-based models are Siegel (1974), Allen (1978), Selkirk (1982), Kiparsky (1982c), Mohanan (1986), and Giegerich (1999).

Stratum-models traditionally assume that the specifications determining the stratum membership, and thus the phonological and morphological processes, are associated with specific affixes. Giegerich (1999), however, proposes that the strata originate from the properties that are associated with bases.

Level-ordering elegantly captured a range of restrictions that characterize English derivational affixation and appeared to significantly improve predictions of possible and impossible combinations. Yet, level-ordering has been criticized as theoretically weak and lacking predictive power (e.g. Fabb 1988, Plag 1996, 1999, Plag and Baayen 2009). Plag demonstrates that there are very many ordering restrictions within each level which are not accounted for by level-ordering, and that the level-ordering hypothesis

has many counter-examples. There are some cases where level 2 affixes appear to attach before level 1 affixes and many constraints on the relative ordering of affixes within each level. There are also so-called Bracketing Paradoxes to be dealt with—cases such as *un-grammatical-ity*, where Level 1 suffix *-ity* must attach to the base containing the Level 2 prefix *un-* to satisfy categorical selectional restrictions (see Fabb 1988). Often suffixes cannot be straightforwardly assigned to a level because in different words they appear to show properties which belong to different strata. As outlined above, many bases have their own selectional restrictions which are in any case needed to account for the stacking behavior within each level. And once these are taken into account, Plag (1999: 92) argues, there is little extra explanatory power offered by a level ordering theory. He concludes that the additional elements needed in the model to avoid wrong predictions considerably weaken the whole level-ordering approach.

While a strict level-ordered approach may not be desirable or necessary, the affix-ordering generalization was clearly still an attempt to deal with a real phenomenon. There is an overall tendency, in many languages, for affixes which are in some sense “closer” to the stem (semantically, syntactically, and/or phonologically) to appear closer to the stem than affixes which are more distant.

21.6 FIXED ORDER AND FREE VARIATION

As has been discussed, some pairs of affixes in some languages can vary in the order in which they occur, with consequences for semantic scope. Examples from Yup'ik and Kwakw'ala were given in (1) and (2). In the absence of scopal restrictions, it is usual for affixes to occur in a fixed order. The level-ordering approach outlined in Section 21.5, of course, would predict that there would be many cases where we get BASE-AB, but not BASE-BA, if A attached at level 1, and B at level 2. But even within so-called levels, if one order of a pair of affixes exists, the other usually does not. However, despite the fact that a fixed order is frequent, it is by no means a universal.

As exceptions to fixed order, we find languages with affixes which occur in multiple orders, with no apparent consequences for meaning. Ryan (2010) for example, discusses a case from Tagalog in detail. In this case, a reduplicant affix marking the contemplated aspect can occur before or after the morpheme over which it holds scope. An example is given in (4).

- (4) a. *pà:* -pa-buks-án 'RED-CAUS-open-LT'
 b. *pa- bú:* -buks-án 'CAUS-RED-open-LT'
 'will cause to open'

(Ryan 2010: 763)

Ryan's account of affix-ordering relies on gradient bigram morphotactics. A learner acquires knowledge about the likelihood of pairs of affixes co-occurring in a particular

order. This is related to accounts which rely on “selectional restrictions.” However, it does not particularly privilege the role of the affix or the base in driving the restrictions. In the case of Tagalog, Ryan shows that the pattern of variability in Tagalog is compatible with learners learning the co-occurrence restrictions in a systematic manner and extending them by analogy to create new forms.

Such an account can deal with variable order well (unlike many accounts of affix order). However it does not, in itself, explain why such variability is not frequently observed across languages. Despite the fact that some languages contain affixes that occur in both orders with respect to one another, there is a strong tendency toward systems with a more predetermined order. Ryan argues that the cross-linguistic tendency toward a fixed order relates to the diachronic patterns in which units tend to get morphologized. These privilege the emergence of fixed order systems.

21.7 ACYCLICITY

Some studies have looked at the degree to which restrictions on orderings are transitive. If we have BASE-AB and BASE-BC, for example, does it follow that BASE-AC is attested? And do the prohibited forms *BASE-XY and *BASE-YZ, imply the ungrammaticality of *BASE-XZ? This general question has been investigated for English suffixes by Hay and Plag (2004) and Plag and Baayen (2009). What they show is that, at least for English, affixes are organized remarkably acyclically. Hay and Plag, for example, show that it is possible to organize English derivational affixes in a hierarchy, such that affixes can attach to something above it in the hierarchy, but not below it. An ordering showing this is given in (5).

- (5) (1) th
 (2) en
 (3) er/ee
 (4) ling/ess
 (5) ly/dom
 (6) ful (N)
 (7) ish/hood/ship
 (8) less
 (9) ness

There are very many attested combinations where an affix listed with a higher number attaches to an affix listed with a lower number (e.g. *-th-en*, *-er-ship*, *-ess-less*). Affix combinations in the other order are very unlikely, except in the case of systematic word-based exceptions.

This is shown for an expanded set of suffixes in Plag and Baayen. There are very few exceptions, and the chances of this happening by chance are exceedingly small. The

acyclic pattern has also been found with English prefixes (Zirkel 2010) and compounds (Baayen 2010b)

Some follow-up work in other languages has also found evidence for such acyclicity—revealing a strong tendency toward transitivity in ordering relationships. Such a hierarchy can be created for Italian (Gaeta 2005, Talamo 2011), Russian (Parker and Sims 2012), Dutch (Baayen and Plag 2007), and German (Hilkenbach 2011).

While transitivity of ordering relationships appears to be quite common, there are certainly exceptions to this pattern. Ryan (2010) for example, summarizes literature illustrating cases of non-transitivity from Quechua, Tucano, Wolof, and Navajo.

In doubly derived Chumbivilcas Quechua verbs, for instance, *ri* “inchoative” can only precede *sch*i “assistive” and *sch*i can only precede *na* “reciprocal” (Muysken 1988: 263). If transitivity held, *ri* would have to precede *na*. But, in fact, *na-ri* is the only acceptable order, violating transitivity. (Ryan 2010: 780)

Clearly, acyclicity is not universal but why should it occur at all? As argued by both Hay and Plag (2004) and Plag and Baayen (2009), the chances of such an acyclic pattern emerging by chance are vanishingly small.

Analyzing data from English suffixation, Hay and Plag argue that the English pattern is a result of “Complexity-Based Ordering.” They originally found the acyclic pattern by hypothesizing that it should follow from a theory of affix ordering that invokes the parsability of the base. Hay (2002, 2003) argues that many of the selectional restrictions observed by Fabb (1988) are motivated by psycholinguistic principles, and can be reduced to a more general principle which constrains affix ordering, while still capturing the spirit of the original level-ordering approach. For example, one of Fabb’s observations is that many suffixes cannot attach to already suffixed words. Hay (2002, 2003) claims that this is one extreme of a continuum of affix behavior, in which affixes are sensitive to the amount of internal structure of the word to which they attach. More particularly, she claims that affixes cannot attach to words which contain a stronger boundary than they themselves create. The central idea is that “*an affix that can be easily parsed out should not occur inside an affix that cannot*” (2002: 528). This account was dubbed Complexity Based Ordering by Plag (2002).

So-called Level 2 affixes are transparently attached and so inherently more parsable and separable than Level 1 affixes. So this generalization captures most of the phenomena that level-ordering was designed to explain and retains its overall spirit. However, it does not posit any particular division between two levels, but rather posits a continuum of separability and a constraint against more separable affixes occurring inside less separable affixes.

Some affixes within complex words are readily isolatable while others are not. For example, *-ish* in *grayish* appears more easily separable from the base than *-ic* in *scenic*. Some affixes tolerate no internal structure while other affixes tolerate it to some degree (Hay 2002: 527). The overall effect is that the less transparent an affix is, the more it resists attaching to already affixed words. Even the same suffix in different words shows

different degrees of parsability (e.g. *-ment* in *government* vs. *discernment*). This leads different words containing the same affix to differ in their openness to further affixation (cf. *governmental*, vs. **discernmental*). Complex words therefore can be complex to a varying degree, depending on how strong a boundary an affix creates between it and the base. Central to Hay's model is the notion that the strength of the morpheme boundary is gradient and so is the degree to which words are decomposable in speech processing.

If an affix that can be easily parsed out should not occur inside an affix that cannot, then this would predict that it should be possible to arrange affixes on a hierarchy of parsability, argue Hay and Plag (2004). As outlined above, they show that English derivational affixes can be arranged on a hierarchy which reflects ordering restrictions, in that any affix that follows the affix X on the hierarchy can attach to a word already affixed in X, whereas any affix that precedes the affix X on the hierarchy resists attaching to a word already affixed in X. The hierarchy provides an account for why affix orders as in *atomlessness* are perfectly acceptable but ones like in **atomlessity* are not (the boundary created by *-ity* is weaker than that created by *-less*). Central to the account though, is that parsability varies in a word-specific manner. Variable decomposability of words predicts variable word-specific behavior. An affix may resist attachment to a complex word that is highly decomposable but may attach to a complex word that is less decomposable. For example, *government* is less decomposable (has less internal structure) than *bafflement* or *discernment* causing *governmental* to be a better word than *bafflemental* or *discernmental*. Thus, affixes are sensitive to the internal complexity they create within a multimorphemic word, and there are systematic word-based exceptions to the hierarchy shown in (5).

Why then is *government* less decomposable than *bafflement*, or *discernment*? Hay's position is that decomposability is largely driven by psycholinguistic speech perception processes, and the processes underpinning word recognition. These include boundary phonotactics, which listeners use to locate word-boundaries (McQueen 1998, van der Lugt 1999), and frequency relationships between the base and the derived words.

Government, for example, is quite frequent and much more frequent than *govern*. *Discernment* is infrequent and much less frequent than *discern*. People tend to rate the latter as more easily decomposable than the former (Hay 2002). This different level of decomposability is important in determining what affixes can then follow the word. The suffix *-al* sits much more comfortably in *governmental* than *discernmental*. Hay (2002) argues that this is an effect of the different degrees of decomposability of the base words.

A predicted consequence of Complexity-based Ordering is that one should be able to order affixes on an overall hierarchy of juncture strength, and that this should reflect the order in which affixes can occur. There should be local exceptions to the hierarchy in cases when an affix is not particularly parsable from particular base words, and these are attested. Overall, an acyclic pattern seems to hold. In English, argue Hay and Plag (2004), the pattern is consistent with Complexity-Based Ordering in that it seems to relate to parsability; the order in which affixes occur in the hierarchy

correlates with measures relating to their parsability and productivity. More productive affixes occur later in the word (and so further down the graph in (5) (Hay and Plag 2004).

However, other work investigating this question in other languages has found the presence of acyclicity without any clear correlation with parsing or productivity factors (Baayen and Plag 2007, Gaeta 2008, Hilkenbach 2011). And a later expansion of Hay and Plag's (2004) data set to a larger set of English suffixes shows that the correlation with affix productivity is robust but the correlation with parsing is not (Plag and Baayen 2009). Plag and Baayen examine word naming latencies and lexical decision data and show that the median reaction time to words containing a particular affix relates, in a non-linear fashion, to rank in the ordering hierarchy. Reaction was fastest to words containing affixes at the beginning and end of the hierarchy. They argue that this is because affixes which tend to occur close to the stem "enjoy the advantages of storage," whereas those that tend to occur far from the stem "enjoy the advantages of efficient parsing" (Plag and Baayen 2009: 138).

Noting that the evidence of acyclicity is much stronger in English than its correlation with productivity, Plag and Baayen also speculate about other reasons why we might see the acyclic pattern. This question is increasingly apt with the report of languages that show acyclicity with no evidence for relationship with productivity (Dutch: Baayen and Plag 2007; Italian: Gaeta 2008, German: Hilkenbach 2011). Plag and Baayen suggest that this may be related to optimization for the listener, in terms of their ability to predict what might come next. In a transitive, acyclic system, the conditional probability of each affix given the previous one is maximized. "In short, acyclic graphs afford enhanced anticipation (in comprehension) and enhanced planning (in production) compared to graphs with many cycles" (Plag and Baayen 2009: 145).

As pointed out by Ryan (2010), transitivity of restrictions and a fixed order for affixes are the norm. He argues that the reasons behind this norm are likely driven by diachronic patterns and by functional factors, such as those discussed above. However, he argues that the grammar can simply be modeled as a series of weighted bigram constraints. Such a grammar can also capture those systems which allow variability and non-transitivity.

21.8 RECURSION

One particular exception to the notion of a fully acyclic system for affixation is recursion.

Recursion refers to the repeated application of a process to an existing structure that was obtained using the same process. The term is more commonly encountered relating to syntax and compounding than to derivational morphology. Recursion in derivational morphology means the application of the same affix to a base that already contains that affix. Although affixation is generally considered non-recursive (e.g. Uhlenbeck 1962,

Anderson (1985; also in Štekauer et al. 2012) shows that the Kwakw'ala suffixes *-exsd* 'want' and *-amas* 'cause' (discussed in example (2) above) create a possibility for recursion between the two verbal suffixes when attached to the verb *q'aq'óla* 'learn':

(8)

q'aq'ola	'learn'
q'aq'olaamas	'cause to learn, teach'
q'aq'olamadzexsd	'want to teach'

The same suffixes appear in opposite orders corresponding to different meanings “cause to want” and “want to cause.” Because the verb “teach” itself is a causative of “learn,” new forms of arbitrary complexity can be produced: “want to cause to want to cause to . . .”

If all phonological, morphological, syntactic, and semantic constraints which regulate affixation are fulfilled, there is no limitation to the number of derivatives the loops produce. However, usage shows that derivational loops do not apply indefinitely for pragmatic reasons (Lieber 2004: 167–70). Each derivational cycle must add new useful semantic content to the derivative. Intensifying prefixes (or suffixes) can apply in succession as long as the intensifying function remains useful. Transpositional (category changing) suffixes, which have little semantic meaning, are subject to the same usefulness prerequisite. Although in theory the loops can attach new derivational layers to the base, complex derivatives soon stop being informative and become difficult to interpret. Transpositional loops sustain their interpretability slightly longer if the first derivative has lexicalized and detached semantically from its form (cf. *business*, *emergency*). Besides pragmatic reasons, (morpho-)phonological constraints, such as avoidance of placing similar or identical sounding syllables next to one another (haplology), may also limit recursion (Plag 1999: 47). Thus, even though derivational recursion does exist, it is much more restricted than syntactic recursion.

21.9 CONCLUSION

The problem of affix-ordering has been, and continues to be, a domain in which there is extensive investigation and theoretical work. While there are many cross-linguistic tendencies in the factors that influence the ordering of derivational affixes, there is also considerable cross-linguistic variation. It is thus a fascinating area for understanding both cross-linguistic tendencies and cross-linguistic variety. The common tendencies reveal important aspects about processing, storage, and word-structure formation strategies. And the rich variety across languages provide insight into different ways of being “optimal,” and the range of solutions to the problem of organizing a particular inventory of affixes in a particular linguistic context.

ACKNOWLEDGMENTS

The publication of this chapter was made possible through the support of a Subaward under a grant to Northwestern University from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. The work was also supported by a Rutherford Discovery Fellowship to the second author. We are grateful to the editors for their valuable feedback.

CHAPTER 22

DERIVATION AND HISTORICAL CHANGE

CAROLA TRIPS

22.1 INTRODUCTION

THIS chapter deals with derivation from a historical perspective. So far, the chapters in this handbook have taken a different perspective and investigated derivation from a synchronic point of view. Thus, questions relating to the development and change in derivation have not been raised so far. In this chapter, the questions that will be addressed are: (1) Where does derivational morphology come from? (2) Are there general developmental stages that must be assumed for all derivational items? (3) What is the trigger for the development of derivational items? (4) Does the lexical semantics of derivational items determine the product of derivation? In the following, we will deal with these questions in more detail.

In the Germanic languages and especially in English and German prefixation and suffixation are the most prominent processes of affixation. Both native prefixes and suffixes have developed out of independent words via a stage where they acted as the first or second part of a compound. In (1a) the Old English (OE) preposition *ofer* ‘beyond, above’ could be preposed to a verb like *climban* ‘climb’ which resulted in a new word-formation *oferclimb(en)* (Middle English, ME) with the meaning ‘climb above/beyond.’ In the course of time, *ofer* developed into a bound morpheme, a prefix, which has to be attached to verbal (and other) bases.

The same process applies to the rise of suffixes. The once free morpheme *hād* ‘state, office, rank’ could function as the second part of a compound as in *child hōd*, and gradually developed into a suffix as in Present-day English (PDE) *childhood*. The rise of prefixes and suffixes is illustrated in (1):

	<i>free morpheme</i>	<i>compound</i>	<i>derivative</i>
a.	OE <i>ofer</i>	ME <i>oferclimb(en)</i>	PDE <i>overclimb</i>
b.	OE <i>hād</i>	ME <i>child hōd</i>	PDE <i>childhood</i>

Another cause for the emergence of affixes in a language is language contact. Prefixes and suffixes may be transferred from their source language to a recipient language ready made, as part of simplexes. As soon as a sufficient number of analysable words with the same affix has been introduced into the recipient language, the structure of these words is decomposed, the affixes are analysed as word-formation elements, and become part of the inventory of the recipient language. We will see below that one such instance of suffixes is PDE *-able* which was transferred from Old French (OF) to Middle English.

A further process giving rise to affixes which will not be discussed here in detail is clipping. Both prefixes and suffixes can develop by this process, for example the prefix *mini-* (clipping from *miniature*) or the suffix *-gate* (clipping from *Watergate*). The latter development includes the reanalysis or reinterpretation of word structure: *-gate* came to be associated with an actual or alleged scandal and since then has been used as a suffix which is added to places where a scandal occurred (*Hollywoodgate*) or to persons implicated in a scandal (*Billygate*). Other instances of reinterpretation are *-burger* (*hamburger*, *cheeseburger*, *turkeyburger*, etc.) and *-teria* the latter of which is based on the foreign word *cafeteria* (*chocolateria*, *drugteria*, *fruiteria*). Further examples are discussed in detail in Marchand (1969).

Finally, to complete the picture, affixes may arise through the reinterpretation of so-called neoclassical coinages. These “compounds” are of Latin and/or Greek origin and can be attributed to the learned and scientific vocabulary which was introduced into English from the 16th century onwards. Although neoclassical elements like *bio-* in *biochemistry* and *biology*, and *-logy* as in *biology* and *geology* could be analyzed (or reinterpreted) as prefixes and suffixes respectively, they do not really behave like affixes because they can combine with bound roots, with words (*bio-diesel*, *garbageology*) or with another combining form. That is why the status of these items is debatable, and some have proposed to call them neoclassical compounds (e.g. Plag 2003; see also Olsen, this volume).

Much has been written about derivation, and more generally about word formation, from a theoretical and synchronic perspective describing and analyzing instances of derivation in many different contemporary languages. Much less attention has been paid to historical aspects of derivation (or word formation), especially to the rise of affixes and systematic descriptions and analyses of affixal inventories available in diachronic stages of languages. In his 2004 paper, Dietz states that although phonology, inflectional morphology, the lexicon, semantics, and syntax are well established traditional fields of research on Old English, word formation still has to catch up considerably. This also applies to other historical stages of English like Middle English and Early Modern English. Even most current work on these stages attribute little space to this field of linguistics.¹ Exceptions are Kastovsky’s (1992b) and Nevalainen’s (1999) seminal contributions to the *Cambridge History of the English Language* (Kastovsky for Old English, Nevalainen for Early Modern English) as well as Marchand’s

¹ See for example the two volumes on English Historical Linguistics edited by Bergs and Brinton in the HSK series in 2012 where marginal descriptions of word formations are found in only two contributions: Kastovsky on morphology (vol. I, part II, ch. 9), and Cowie on morphology in Early Modern English (vol. I, part V, ch. 38).

(1969) and Koziol's (1972) indispensable handbooks of English word formation. Although both handbooks focus on Present-day English, they include information on the development of word-formation processes in general and the origins and development of each lexical entity discussed. *The Oxford Reference Guide to English Morphology* by Bauer et al. (2013) is the most recent comprehensive work on the topic containing voluminous material on both derivation and inflection based on corpus material.

Things are different for other Germanic languages such as German, for example. Here, the study of word formation has a longer tradition although most of the studies written before *Grimm's Grammar* (1826) saw words as atomic entities which had to be defined according to their etymon. Both Grimm and Paul (1880, 1920) applied a holistic approach to the study of word formation in that they investigated the formal and semantic side of lexemes and their development. In his famous *Prinzipien der Sprachgeschichte*, Paul (1880), for example, devotes a whole chapter to the development of word formation and explicitly discusses the rise of prefixes and suffixes. Erben (1964, 1972, 1977) was the first to postulate a description of historical word formation based on structural synchronic and diachronic grounds. Studies dealing either with synchronic descriptions of the periods of German or providing a comprehensive picture are, for example, Henzen (1947), Naumann (1986), Solms (1990), and Splett (1993).²

In the following, the four questions posed above will be addressed. The focus of our attention will be the history of English; comparisons with German will be made where appropriate. Before we will take a closer look at prefixation and suffixation as the two most prominent derivational processes in the history of English, in Section 22.2, descriptions of the derivational systems of Old English, Middle English, and Early Modern English will be provided. The affixes in question will be listed and described in formal and semantic terms. In Section 22.3, we will concentrate on some selected case studies of prefixation and suffixation and discuss some analyses proposed in the literature. Language contact as a possible source for the rise of affixes will also be briefly addressed. As mentioned above, these three case studies might be offered as examples of general historical processes that can lead to the development of or changes in derivation.

22.2 STAGES IN THE HISTORY OF ENGLISH

22.2.1 Affixation in Old English

Old English (c. 800–1150) is an inflecting language with a core Germanic vocabulary and a much more variable word order than Present-day English. It still shows many genuine Germanic features and thus resembles other (West) Germanic languages like New

² For studies on the development of Romance affixes, see Darmesteter (1877), Zwanenburg (1983), Fleischman (1977), Grossmann and Rainer (2004), and Lütke (2005).

High German (NHG) much more than the later stages of English, that is Middle English (c. 1150–1500) and Early Modern English (c. 1500–1700). One reason for this development is certainly the fact that both ME and EModE were affected by foreign influence to a higher degree on the level of the lexicon. Old English nominals inflected for case, number, and gender, verbs inflected for person, number, tense, and mood. Since morphological inflection indicated grammatical specifications, the arguments of verbs could occur in a number of different positions.

Another genuine Germanic property is *ablaut* (stem vowel change) not only occurring in the verbal inflection of strong verbs but also in products of derivation. A strong verb like *drincan* (with the forms *dranc*, *druncon*, *gedruncon*) ‘drink’ built derivations like *drinc* ‘drink, drinking,’ *drincere* ‘drinker’ with the suffix *-ere*, *druncennis* ‘drunkenness’ with the suffix *-nis*, but also forms like *drenchus* ‘drinking house’ which derives from the weak verb *drencan* built by the process of *umlaut*, also called *i-mutation* (compare NHG *trinken*, *Trinker*, *Trunkenheit*, *Trinkhaus*, *tränken*).

Other morphophonemic alternations occur as (a) consonant gemination (accompanied by *i-mutation*): *secgan* ‘say’ (V) → *sagu* ‘saying’ (N); (b) palatalization/assibilization: *cēosan*³ ‘choose’ (V)/*curon* (preterite plural) → *cyre* ‘choice’ (N); (c) change from /æ/ to /a/: *faran* ‘go’ (V) → *fær* ‘journey’ (N). Some prefixes exhibit stress-conditioned alternations which results in differences between verbs and nouns with the same prefix. In verbs, prefixes are either unstressed or bear secondary stress, in nouns they bear primary stress. Thus, we find alternations like *begán* ‘go over, worship’ (V) → *bígèng* ‘worship’ (N) which are often accompanied by allomorphy. The source of this type of alternation can be found in Proto-Germanic times when word stress came to be fixed on the first syllable (primary stress on the initial syllable). Since prefixed nouns already existed at that time no change occurred, whereas prefixed verbs are a later development and kept primary stress on the stem (for details see Kastovsky 1992b). From these examples, it is evident that stem variability, that is, stem allomorphy, present in inflection and word formation characterized the Old English morphological system, and this is what we expect to find when we take a closer look at prefixation and suffixation.

In Old English, the set of prefixes was quite large and they were widely used with nouns, adjectives, and verbs. Generally, two types of prefixes are distinguished: (a) the primary prefixes, native prefixes, which in Proto-Germanic times developed from free morphemes into bound morphemes via a stage where they acted as the first part in compounds—in OE, these prefixes are fully lexicalized, that is, no longer formally and/or semantically transparent; (b) secondary prefixes follow the same developmental path but arose at later stages in the individual Germanic languages (for a discussion on the classification of prefixes, cf. Dietz 2004).

Attached to nouns, prefixes had the function of deriving adjectives (*drēam* ‘joy’ (N) → *sindrēam* ‘everlasting joy’ (N)), attached to adjectives and verbs they functioned

³ Since in OE the quantity of vowels is phonemic, vowel length is indicated by a diacritic (e.g. *ā*, *ō*) in all of the examples.

as adverbs (*ceald* ‘cold’ (A) → *sinceald* ‘perpetually cold’ (A), *cweþan* ‘speak’ (V) → *mis-cweþan* ‘speak ill’ (V)). All Germanic prefixes developed from adverbs, prepositions, and particles with a local-deictic meaning. OE verbal prefixes modify the lexical meaning of the verb in terms of the prototypical binary features [static] vs. [dynamic], [local] vs. [temporal], [dimensional] vs. [relational]. Moreover, they can characterize verbal aspect, for example *onhætan* ‘inflamm’ (ingressive), *ablācian* ‘grow faint’ (inchoative), *bedrincan* ‘drink up’ (perfective) (cf. Dietz 2004).

In the following, some of the most frequently occurring and regularly attested prefixes are briefly discussed, for a comprehensive account see Kastovsky (1992b) and Dietz (2004). The prefix *be-* is mainly used with verbs adding a number of different meanings to them: ornative *begyrðan* ‘to girdle,’ privative *behēafðian* ‘behead,’ transitivization *beflēogan* ‘fly upon,’ intensification *bebreccan* ‘break to pieces’ (for further details, see Section 22.3.1). Due to the allomorphic nature of OE, we find different variants of prefixes, for example *bi-* which is the stressed variant of *be-*, occurring with nouns as in *bigyrðel* ‘girdle, belt’ adding the meaning “around” or with deverbal nouns as in *bigeng* ‘worship.’ The most frequent prefix in OE is *ge-*. Two functions can be identified depending on whether it is prefixed to nouns or verbs. With nouns it either adds the meaning “collectivity” as in *gefylce* ‘collection of people, army’ or “associativity” as in *gebedda* ‘one who lies in bed with another.’ Prefixed to verbs, *ge-* denotes “perfectivity” or “result” as in *geāscian* ‘learn by asking.’ Sometimes these meanings are no longer transparent as in *gestandan* ‘endure, last,’ or there does not seem to be a difference in meaning between the simplex with or without the prefix as in (*ge*)*campian* ‘strive, fight.’ *un-* is a very frequent prefix and denotes negativity: “not, opposite.” It is prefixed to (deadjectival) nouns (*unlif* ‘death’), adjectives (*unbrād* ‘narrow’) and verbs. Prefixed to verbs *un-* denotes “reversal of a result” as for example in *unbindan* ‘unbind, loosen.’

As noted above, the allomorphic nature of OE is reflected in the OE prefixal system. Moreover, many prefixations often were no longer transparent, and prefixes did not seem to have a significant semantic effect. This might have led to the weakening of the prefixal system, and it is not surprising that at the end of the OE period, the prefixal system was already in a state of advanced decay (Horgan 1980, Hiltunen 1983). Kastovsky notes that the OE system “... was an easy victim both for the Romance invasion of the lexicon and the rise of the postparticle (phrasal) verbs” (1992b: 377). If the OE prefixal system is compared with NHG, we see that many of the Germanic or native prefixes have been retained in the latter language, and are more or less frequently used depending, amongst other factors, on the lexemes they attach to. For example, the prefixes *miß-* (OE *mis-*), *un-* (OE *un-*), and *ur-* (OE *or-*) are attached to nouns (*Mißernte* ‘failure of crops,’ *Urmensch* ‘prehistoric man,’ *Untiefe* ‘shallows’) and adjectives (*mißgelaunt* ‘ill-humoured,’ *urgemütlich* ‘very cosy,’ *unklug* ‘unwise’). The prefix *ge-* (OE *ge-*) occurs with nouns and adds the meaning “collectivity” to the prefixation (*Gebüsch* ‘bushes’). With deverbal nouns the productive model “*Ge* + verbal base + *e*” is built forming Nomen actionis like *Gebelle* ‘incessant barking’ with an iterative meaning. The largest set of native prefixes occurs with verbs, where most of the prefixations are productive. The most productive prefix is *be-* as in *besteigen* ‘climb up’ with many shades of meaning.

The prefixes *ent-* (*entladen* ‘unload’), *er-* (*erfassen* ‘understand’), *ver-* (*verfolgen* ‘follow’), and *zer-* (*zerdrücken* ‘crush’) which in OE were already almost completely lost (residual cases are *fer-rædan* ‘condemn’ and *tor-cyrre* ‘hard to be converted,’ cf. Dietz 2004: 568f) are still productively used to express aspect (for a full and systematic account of NHG prefixes, see Fleischer and Barz 1995). This applies to other Modern Germanic languages such as Dutch or Swedish as well.

While the OE prefixes exhibit allomorphic alternations, OE suffixes may cause morphophonemic alternations in the lexemes they are attached to. The most prominent of these alternations is due to i-mutation (see above) where the vowel of the syllable bearing main stress is assimilated to the following secondary stressed or unstressed vowel. The result is palatalization of back vowels, and palatalization and raising of front vowels. A good example is the suffixation with OE *-ig* deriving denominal, deadjectival, and deverbal adjectives with the meaning ‘characterized by, having.’ The products of suffixation show i-mutation: *stān* (‘stone’)—*stænig* (‘stony’), *gesund* (‘sound’)—*gesyndig* (‘sound’). In the course of the OE period alternations of this type became less and less frequent and are only sporadically found.

In contrast to a set of prefixes moderate in size, suffixation in OE displays a considerable number of suffixes which can be attached to nouns, adjectives, and verbs. In the following, some of the most productive nominal, adjectival, and verbal suffixes are provided (based on Kastovsky 1992b). In Section 22.3, we will deal with some of these affixes in more depth.

To derive nouns, the following suffixes can be used in OE: *-dōm* (*cāserdōm* ‘empire’) deriving nouns with the meaning ‘state, condition, action of,’ *-ere* (*leornere* ‘disciple’) deriving deverbal agent nouns, *-hād* (*cildhād* ‘childhood’) deriving abstract nouns denoting ‘state, rank, order of N,’ *-scipe* (*freondscipe* ‘friendship’) deriving masculine abstract nouns denoting ‘state, act, fact,’ and *-ung* (*huntung* ‘hunting’) deriving deverbal nouns with the meanings ‘action, agent, object, instrument, location.’ One suffix that is particularly productive is *-ness*. Deverbal nouns are derived from either the infinitive or from present and past participles: *costness* ‘temptation,’ *āstandendness* ‘continuance,’ *upārīsenness* ‘resurrection,’ *foreset(ed)ness* ‘purpose.’ Accordingly, doublets can be found with no meaning difference: *forgifness/forgifenness* ‘forgiveness,’ *ongitness/ongitenness* ‘understanding.’ As with the other deverbal action nouns discussed above, a range of meanings occur, from action to object/result, instrument, and location.

In OE, adjectives can be built with as large a choice of suffixes as nouns. The suffix *-bære* which derives from the verb *beran* ‘carry, bring’ is frequently used to attach to nouns meaning ‘carrying, having, productive of.’ Whereas the suffix was lost in ME times, in other Germanic languages like German it is used to build adjectives from transitive verbs (*tragbar* ‘portable/acceptable,’ *trinkbar* ‘drinkable’) in a highly productive way. In Section 22.3.3 we will discuss the loss of this suffix in the context of language contact with French.

Another suffix frequently used to build new lexemes is *-ig*. It can be attached to nouns, adjectives, and verbs and adds the meaning ‘characterized by, having.’ Both

forms with and without i-mutation are found, examples of denominal, deadjectival, and deverbal adjectives are *blissig* 'joyful,' *gesyndic* 'sound,' and *cwānig* 'sad, sorrowful.' According to Kastovsky (1992b), the suffix *-lic* has the same meaning and attaches to the same stems, and that is why many doublets are found. Some of the many examples of derivations with *-lic* are *andgieticlic* 'intelligible,' *dēoplic* 'deep,' (*ge*) *cwemlic* 'pleasing, satisfying.'

In the preceding paragraphs, the considerably large set of OE verbal prefixes was described, most members of which were productively used. In contrast, only a few derivational verbal suffixes exist which do not exhibit high productivity. One of them is *-s(ian)* (*cildsian* 'be childish') building denominal and deadjectival verbs as *metsian* 'feed, furnish with provisions' and *clænsian* 'make clean.'

The nature of the formation of adverbs reflects the OE system as an inflectional system with stem-based morphology. Since the formation of adverbs in OE times does not add meaning to the deadjectival and denominal adverb, it has been attributed to inflection. Others have defined it as a derivational process as it involves a change in word category. So it seems to be a borderline case between inflection and word formation.

One of the few suffixes available to form adverbs is *-e* (*dēope* 'deeply'), which shows by far the highest frequency. Since some adjectives also end in *-e*, adjectives and adverbs can be homophonous, for example *milde* 'merciful(ly)'. Some adjectives are built with the suffix *-lic*, and since *-e* can also be attached to these forms (*sārlic* 'grievous' → *sārlice* 'grievously'), *-lice* was secreted to directly form adverbs from adjectives without the *-lic* suffix (*blindlice* 'blindly'). Denominal adverbs can be formed in the same way (*freond* 'friend' → *freondlice* 'friendly').

The suffixes *-es* and *-um*, also used to derive adverbs, are inflectional endings, the former marking masculine, genitive, singular, the latter marking dative plural.

Due to the loss of the OE inflectional system and to the increasing opacity of alternating derivation, by the end of the OE period the morphophonemic system was restructured. Thus, OE can be seen as a transitional stage where the system changed from stem-based to word-based morphology, and some results of this change can already be seen in the ME system which we will turn to in the following section.

22.2.2 Affixation in Middle English

It is a well-known fact that in ME many of the nominal and verbal inflections were drastically reduced, for example the different grammatical forms of the OE nouns *stān* 'stone' were reduced to three: *stān*, *stānes*, *stāne*. Nevertheless, residues of the system still existed, which is why at that time derivation may be word-based or stem-based. As long as traces of stem-based morphology exist, like for example infinitival endings as in *declaren*, a derivation like *deklaracioun* with a non-native suffix still is a stem-based derivative. The strong tendency towards word-based morphology, however, cannot be denied and by the end of the ME period hardly any traces are left of the OE system.

Prefixation as a means to build new lexemes is on the decline. The observation that some of these prefixes were no longer productive and did not add further meaning to prefixations might have played as much of a role as the massive borrowing of lexemes from Old French. Although some of the prefixes discussed above are still found in native lexemes inherited from OE, they are no longer used for building new lexemes. For example, the prefix *ge-* which occurs in ME as *i-/y-* (*ydrunken* 'drunk') is still used to mark the old past participle form of verbs in the southern dialects. Others were no longer productive and rapidly disappeared altogether. The only prefix which was retained is *un-* but it could no longer denote "intensity" nor "pejoration" being added to nouns and adjectives. Instead, it became more and more productive with verbal bases adding a reversative meaning by the derived lexeme (*unclose* 'to open,' see also above).⁴

The first OF prefixes coming into the English language are *de-/des-* (Latin *dis-*) as in ME *desarmen* 'divest of arms of armor,' *despisen* 'to look down upon' (cf. MED⁵) adding an adversative or privative meaning, and *mes-* which fell together with OE *mis-* (ME *misgiltēn* 'to do wrong')⁶ In the course of time, the density of French loan words increased with the rate of new lexemes reaching a peak in the second part of the fourteenth century (cf. Burnley 1992: 431). In the fourteenth and fifteenth centuries more borrowings were made, especially abstract nouns which contained the prefixes *con-* (*confermen* 'endorse'; the meaning is not entirely clear at that time), *ex-* (*expounen* 'set forth') "out, out of," *pre-* (*prescriben* 'order') "before," *pro-* (*prolongen* 'to lengthen in time') "in favour of" and *trans-* (*transcriben* 'copy out') "across, over." Overall, most of the native prefixes vanished and were "replaced" by non-native ones, that is, prefixes of OF and Latin origin.

Burnley distinguishes three successive stages in the process of adopting and assimilating non-native affixes. In the first stage, an affix is part of a lexeme which is borrowed as such (a simplex) and assimilated into the lexicon (and grammar) of the recipient language. In the second stage, speakers of the recipient language analyse the word structure of these loans which results in the formation of new lexemes with these affixes with a restricted set of bases. In the third stage, the affixes are no longer felt to be foreign and are widely and productively used with bases of native and non-native origin. It is at this point that affixes have become part of the recipient language's derivational system. Obviously, the prefixes discussed above reached these three stages to different degrees, and this is also what we expect to find for non-native suffixes.

In the inherited suffixal system, changes were not as dramatic. Of the OE suffixes, more than three-quarters were retained in ME, for example *-dom* (*dukdom* 'duchy'), *-ere* (*clothere* 'cloth maker'), *-hede* (*kinghede* 'kingdom'), *-ness* (*cherishness* 'affection'), *-ship* (*sadship* 'maturity, seriousness'), and *-ung/ing* (*deoring* 'darling') building abstract or

⁴ According to Burnley (1992: 447) OE *un-* was probably confused with OF and Latin *in-*. It seems to appear especially with adjectives formed with the suffix *-able*.

⁵ For the ME prefixations and suffixation listed in the paper, the Middle English Dictionary (MED) online (<<http://quod.lib.umich.edu/m/med/>>) was consulted.

⁶ Due to the many spelling variants of ME, I used the headword forms of the affixes given in Burnley (1992), Dalton-Puffer (1996a), and the MED.

common nouns, the suffixes *-ig* (*holie* ‘divine’), and *-ish* (*foxish* building adjectives, and the suffix *-sian* (*clensen* ‘cleanse, purify’) building verbs (cf. Burnley (1992) and especially Dalton-Puffer (1996a) for a full discussion and a full list of suffixes retained).

Although many of the native suffixes are still found in ME, not only as part of simplexes but also as items productively building new lexemes, OF and Latin were prolific sources of non-native derivational suffixes. The set of native suffixes building abstract or common nouns described above was enriched by *-acio(u)n* (*vexacioun* ‘physical suffering’), *-erie* (*husbondrie* ‘housekeeping’), *-ite* (*enemite* ‘rivalry, hostility’), and many others.

The native suffixal system building adjectives was enriched by suffixes of Romance origin like *-able* (*knouable* ‘capable of being known or learned’). The most frequent adjectival suffixes in terms of type frequency are *-ly*, followed by *-ful*, *-able*, *-ward*, and *-ous* (cf. Dalton-Puffer 1996a).

It was shown above that the set of native verbal suffixes was very small and that it did not productively build new lexemes. The verbal suffixes of Romance origin did not productively build new lexemes either. The suffix *-ify* (*certify*, *magnify*) exhibits the highest number of occurrences in Dalton-Puffer’s study, and may thus be regarded as a suffix in the ME period (for the other suffixes see the *Oxford English Dictionary* (OED) and Marchand 1969).

All of the suffixes predominantly borrowed from Latin and/or OF underwent the process of adoption and assimilation defined by Burnley. Some of them never got beyond stage one, some of them reached the third stage which resulted in the full integration of these “derivational devices” into the native (English) system. In Dalton-Puffer’s study, most of the Romance suffixes start to show greater numbers of (analysable) derivations in the ME₃ period which is dated 1350–1420. This implies that at that time the suffixes were established as word-formation elements, they had reached the third stage and were completely assimilated to the Germanic system.

By the end of the ME period, the morphological system was predominantly word-based, and the derivational system had been enriched by a considerable number of affixes of Romance origin. This is the basis of the Early Modern English (EmodE) period. What is striking in this period is the vast expansion of the system productively building new lexemes. In the following, we will see that in that period foreign influence (in more general terms) is reflected in word formation by a sharp increase of non-native elements as productive affixes.

22.2.3 Affixation in Early Modern English

The Early Modern English period has been said to be the period with the fastest vocabulary growth in the history of English. Based on statistics from chronological dictionaries like the *Chronological English Dictionary* (CED), the time span between the early 16th to the mid-17th century has been identified as the peak of massive lexical borrowing (cf. Burnley 1992). But it is not only the borrowing of foreign words that contributed to the

expansion of the lexicon but also the productive use of word-formation processes. To be more precise, quite a number of the OF affixes that had come into English during the ME period as part of simplexes, are now used as transparent word-formation elements productively building new lexemes.

In the 16th century, the English vernacular was felt to be insufficient for translations of the Bible and the classics because it was said to be inexpressive and ineloquent by quite a number of learned people. This led to debates between neologizers, etymologists, and purists which gradually gave rise to the Inkhorn Debate. The neologizers argued in favor of expanding the lexicon by adding foreign words from the classical languages. This is why most lexemes were borrowed from Latin. Between 1560 and 1670 more than half of the loan words which are attested in the *CED* are Latin loans. These loans were specialist terms from the fields of medicine, zoology, botany, and theology. These newly adopted affixes with Romance and classical bases clearly outnumbered the native ones both in type and token frequencies. Most of them had the status of adoptive affixes because they built new lexemes with native bases.

According to Wermser (1976), at the beginning of the EModE period the native affixes amounted to about 80% of all new derivations, whereas at the end of this period the amount had decreased to about 30%. Thus, the situation completely changed, especially if we take a look at the development of the prefixal system. Whereas between the OE and ME period quite a number of native prefixes were lost and only some of Romance origin were borrowed, in EModE prefixes of Romance origin were massively borrowed and introduced into the EModE lexicon. These prefixes contribute to many semantic categories expressing negative, locative, temporal, opposing, pejorative, intensifying, and quantitative meanings. In the following, some examples are given.

One of the most common negative prefixes is *un-*. Whereas in ME it almost exclusively combined with verbs adding a reversative meaning, in this period it combines with all types of adjectives (*unfit*, *unearthly*), participial adjectives (*unbecoming*), and since the 16th century with past participles of prepositional verbs (*unheard-of*). It further builds reversative verbs like *undress*, *undraw*, and due to the increasing productivity of denominal conversion verbs, privative verbs with the meanings “objective” (*uncloak*) and “ablative” (*unballast*) are productively formed. It very often combines with verbs derived with the suffixes *-ize* and *-ify* from about the 16th century onwards (e.g. *unbarbarize*, *undeify*).

One of the locative prefixes emerging in the EModE period is the prefix *trans-*. It adds the meaning “across” and is frequently found in Latin loans in EModE times. It frequently combines with denominal verbs (*transnature*), nouns (*translocation*), and sometimes also with adjectives (*translunary*). The only native prefix productively building formations of this type is *fore-* meaning “in front of, before.” It almost exclusively combines with nouns building formations like *forename*, *forehand*, *forewoman*. It also builds denominal and deverbal formations conveying the temporal notion “before” with native and non-native bases as in *forefeel*, *foremention*, *foremother*, *forenight*.

The two temporal prefixes which are most productive in this period are *pre-* and *re-*. *Pre-* meaning “before” frequently combines with verbs of Latin origin building

formations like *precontract*, *prejudge*, *predeterminate*, *preconcert*. It also builds denominal formations like *predisposition* and *pretaxation*.

Two of the prefixes conveying the notions of opposition and support are *anti-* “against, opposing,” and *counter-* “against.” They gained the widest currency in the course of the EModE period. The former builds denominal and deverbal derivations (e.g. *counterplea*, *counterplot*), the latter denominal and deadjectival derivations (e.g. *antipope*, *anticeremonial*, *antihysterical*).

One example of a prefix with pejorative meaning is *mis-* “wrongly, badly, amiss” which had been introduced into English in OE and ME times. In EModE *mis-* is most productive predominantly building verbs (e.g. *misgive*, *misspell*, *misfire*, *misconduct*).

Non-native prefixes expressing degree and size are for example *arch-* “supreme, highest,” and *proto-* “chief, first,” building denominal formations (*arch-duke*, *protoparents*). *Be-* which is the only native prefix still found in EModE times is commonly used to build denominal and deadjectival verbs as for example *begift*, *becrown*, *becalm*, *bemad*. Since it is exceptional in that it has survived and been productive for so long, the properties of *be-* will be discussed in more detail in Section 22.3.

Apart from the prefixes described above, the EModE lexicon was enriched by a set of quantitative prefixes of which *bi-* “two,” *multi-* “many,” *tri-* “three,” and *uni-* “one” are most productively used. Many of the prefixes described here were competing with other prefixes with a similar meaning. For example Latin *multi-* occurred alongside the Greek prefix *poly-* both meaning “many.”

Although a large number of productive prefixes have been newly introduced into the English language in EModE times, they are evidently outnumbered by suffixes. Many native suffixes, which are interestingly most productive, were retained, and a considerable number of non-native suffixes which arose in the context of loan accommodation in ME times still exist.

The largest group are suffixes building abstract and concrete nouns. One of the suffixes emerging in EModE is the suffix *-ee*. It denotes the indirect object of the verbs from which they are derived. According to Nevalainen “[T]he adoption of the passive benefactive suffix *-ee* in Early Modern English marks the only significant semantic addition” (1999: 396). Although the suffix occurred in ME pairs *donor/donee*, in EModE it gains the “beneficiary” character and occurs with native (*trustee*, *drawee*) and non-native bases (*referee*, *payee*).

The only new suffix occurring in EModE times is *-esque*. It derives adjectives from Romance loans, the first native formations are attested from the 18th century (e.g. *carnavalesque*).

All of the other non-native suffixes having emerged in OE, and especially in ME, have become established as word-formation elements building new lexemes with different degrees of productivity. For example, the French suffix *-ery* (*-ry*) denoting “state/business/behaviour of X” (*barbery*, *smithery*), “collectivity” (*peasantry*, *tenantry*), “place of activity” (*brewery*, *fishery*) productively builds abstract and common nouns in EModE. Other examples are the non-native *-ity* and the native *-ness* suffixes most productively deriving abstract nouns at that time. A non-native suffix deriving adjectives from verbs

is *-able* which also becomes highly productive. In Section 22.3 we will take a closer look at its development and properties.

This section sought to provide an overview of the development of the affixal system in English. It was shown that the system changed from being stem-based (OE) to becoming word-based (ME, EModE). Further, the OE prefixal system was deprived of its plethora of native prefixes in ME times but this loss was then compensated by the addition of a large number of non-native prefixes in EModE times. The amount of native suffixes, however, did not vary to the same degree but was rather stable. Although many suffixes were borrowed as well, the most productive suffixes are the native ones in EModE times.

As stated in the introduction, historical aspects of affixation have been paid little attention and can be seen as a hitherto neglected field. This applies to prefixation even more so than to suffixation. The studies published since the 1990s deal with prefixation in the history of English as such, or with one synchronic stage, or with the properties of one or a small set of prefixes: Yamanouchi (1992), OE *be-* in glosses and the Lambeth Psalter; Ogura (1995), verbal prefixes in OE; Lutz (1997), the development of the English prefix verbs with a focus on sound change; McConchie (2000, 2007), the integration of the negative prefix *dis-* in ME and EModE. Kastovsky (2002) dealing with the development of ornative, locative, ablative, privative, and reversative verbs in English, and Dietz (2004) with formal, functional, and semantic aspects of the OE prefix system, are the two most comprehensive and systematic discussions of OE prefixes. Further aspects of prefixation are discussed in the following papers: Dietz (2005), a comparative study on OE *æ-*, *ō-*, *a-* and their OHG counterparts; Chapman and Skousen (2005), an investigation of morphological change in Skousen's Analogical Modeling based on data from ME and EModE; Petré (2006), *be-* in late OE and Early ME; Petré and Cuyckens (2008), see Section 22.3.1; Minkova (2008), properties of the morphology–prosody interface by using prefixation and stress in OE as a test case; Schröder (2008), the productivity of verbal prefixation in the history of English; and Okada (2010), negative prefixes in the history of English.

Concerning studies on suffixation in English from a historical perspective, since the 1990s, the following papers have been published: Suzuki (1990), OE *-nis*; Peters (1991), a survey of the analyses of OE *-ild*, *-incel*; Peters (2006), OE *-ettan*; Dalton-Puffer (1994), the productivity of Romance suffixes in ME; Dalton-Puffer (1996a), a comprehensive corpus-based study of suffixal derivation in ME with a focus on the impact of Romance suffixes; Miller (1997), the borrowing of word-formation elements in the language contact situation between ME and (Old) French; Zergollern-Miletić (1997); morphological adaptation of suffixes building nouns in English borrowed from French; Delmas (2000), *-able* in ME; Möhlig (2001); OE *-nes*; Terasawa (2001), OE *-ere*; Kaunisto (2001, 2004), the alternation of *-ic* and *-ical* in the history of English. Ciszek has dealt with a number of derivational suffixes in her papers from 2002, (ME *-lich(e)*), 2004 (productivity of *-esse*, *-age*, *-crie*, and *-ment* in ME), 2006 (*-dom* in ME), and 2012 (*-ish* in ME), and in her doctoral dissertation from 2008 (a dynamic analysis of the semantic characteristics, productivity, and dialect distribution of ME *-dom*, *-ship(e)*,

-hede, *-ness*, *-age*, *-erie*, *-ment*). Lloyd (2005, 2006, 2007a,b), productivity and semantic coherence of ME and EModE *-ment*, *-ancel/-ence*, *-ation*, *-age*, and *-al*; Martin Arista (2006), OE *-a*; Ibanez Moreno and Gonzáles Torres (2006), a discussion of criteria that define affixes in OE; Jakubowski (2006), suffixes of Romance origin in Late ME romances; Bauer (2007), a corpus-based study of OE *-fæst*; Kornexl (2008), the history of female nouns in *-ette*; Trips and Stein (2008), ME *-able*, see Section 22.3.3; Palmer (2008), a study of borrowed derivational morphology in Late ME records of London Grocers and Goldsmiths; Palmer (2009), the productivity of *-age*, *-ity*, *-cion*, and *-ment* in the history of English (especially from the 14th to the 16th century); Trips (2009), the rise of the suffixes *-hood*, *-dom*, and *-ship* analyzed in a revised model of Lieber (2004); Baeskow (2010), the history of the suffixes *-ship*, *-dom*, and *-hood*; Nagano (2010), the functional change in the derivational suffix *-ing* in the history of English; Torre Alonso (2011), the origin of affixes, (in)separability, affix loops, and closing affixes in OE; Gardner (2012), OE *-dōm*, *-hād*, *-lāc*, *-nes*, *-ræden*, and *-scipe*; Thompson (2012), the interaction of syncope with the OE *-ig*; Vázquez-López (2012), the development and productivity of nominalizations in EModE.

Without doubt, the author who has contributed most to the field of diachronic word formation (of English) is Dieter Kastovsky. With his invaluable work he has provided us with many insights into the status of OE affixes (and compounding) and the development of derivational affixes, and set the basis for further work on this topic. As noted above, his 1992 paper is one of the most frequently consulted contributions to OE semantics and vocabulary which is part of the *Cambridge History of the English Language* (see Section 22.1). In some of his most recently published papers, he has dealt with prefixation in the history of English (2002, see Section 22.2), with aspects like typological changes in the derivational morphology of English (2006a, b), with surveying word-formation in the 20th century (2006c), or most recently, with semantic loans and loan translations in OE (2010). His textbook *A Historical Morphology of English* will posthumously appear 2014.

22.3. THREE CASE STUDIES ON HISTORICAL ASPECTS OF AFFIXATION

In this section, three studies will be introduced and discussed to provide more in-depth analyses of prefixation and suffixation from a historical perspective. The three studies discussed in the following were chosen because they are current investigations of affixation and thus reflect current work in this field. Moreover, each of these studies illustrates a different way in which derivation can change or come into being.

In addition, all of these studies have made extensive use of corpus data from (the same) diachronic and synchronic corpora and are thus comparable from a methodological point of view. In Section 22.3.1 the study by Petré and Cuyckens (2008) on the prefix

be- is discussed. The authors account for its preservation and productivity by making use of Construction Grammar. Section 22.3.2 deals with the study by Trips (2009) on the rise of the suffixes *-hood*, *-dom*, and *-ship* in the history of English. She claims that suffixes generally cannot be rivals, as is often assumed in the literature, because of their different semantic properties. By looking at the history of these three elements, she shows that the meanings of *hād*, *dōm*, and *scipe* are preserved as salient meanings of the suffixes *-hood*, *-dom*, and *-ship* from which metonymic shifts occurred. That is why these suffixes, and presumably other suffixes as well, are not interchangeable. Her analysis is based on Lieber (2004) which was adapted to historical aspects of affixation. The third study, discussed in Section 22.3.3, exemplifies with the suffix *-able* that language contact led to the enrichment of the English word-formation inventory. Trips and Stein (2008) assume that the integration of *-able* into the ME lexicon must be seen as an instance of structural borrowing. They provide proof for their assumption by comparing the syntactic and semantic properties of word formations with *-able* in OF and ME. They claim that under the influence of French, the native suffixal system was restructured leading to the gradual loss of the suffixes *-lice* and *-bære* (OE forms). The authors further show that in more general terms information about event structure must be included in word-formation rules to fully account for the properties of these (and other, similar) suffixations.

Although the three case studies discuss historical aspects of affixation in English, similar processes can be observed in other languages, as has been noted above for other Germanic languages, and Romance languages (cf. footnote 2). The three studies are therefore meant to be representative of general diachronic processes of affixation.

22.3.1 Case Study 1: Preservation and Semantic Shift

We have seen in Section 22.2 that there is a drastic decline of native prefixes from OE to ME. Although the OE prefix *be-* is still found, it is no longer used to productively build lexemes in ME times. As opposed to the other native prefixes, in EModE *be-* is still frequently used, especially in denominal and deadjectival verbs (*begift*, *bemad*, see Section 22.2.3), so it seems to have survived for a much longer time than the other native prefixes. This is also the motivation for Petré and Cuyckens (2008) to take a closer look at the properties of this prefix and to account for its preservation and productivity. Their account is modeled in Construction Grammar (cf. Goldberg 1995, 2006, Kay and Fillmore 1999) which is based on the idea that the primary unit of grammar is a grammatical construction, or construct, which is a pairing of form and content, regardless of its formal and/or semantic complexity and make-up. Thus, all pairings of form and content are constructs, including phrases, idioms, words, and morphemes. Petré and Cuyckens treat the morphological structure they assume for the *be-* prefixations ([*be-* verb stem]) on a par with argument structure constructions proposed by Goldberg.

For them, these constructs, which they call inseparable complex verbs (ICVs), have the defining characteristics of a grammatical construction: their semantics is not predictable because the meaning of the root verb added to the meaning of the prepositional

cognate of the prefix (*be* ‘by’) does not equal the meaning of the whole construction. Further, the form of the ICV is not compositional because the prefix is not an independent component, it exists only as part of the construct.

Petré and Cuyckens distinguish between two types of ICVs, the *be*- predicative construction, and the *be*- non-predicative construction and claim that the latter survived because of its special semantic properties. The example in (2) shows a predicative prefix construction with the verb *bestelan* ‘to move stealthily,’ example (3) shows a non-predicative prefix construction with the verbs *beridan* ‘ride around’ and *began* ‘surround’:

- (2) *Her hiene BEstael se here [...] to Cippanhamme*
 Here him *be-stole: IND.* the army [...] at Chippenham
PST.3SG
 & geridon Wesseaxna lond
 and occupied West-Saxon’s land
 ‘This year the army **stealthily moved** itself in [...] at Chippenham and occupied the territory of the West-Saxons.’

(Petré and Cuyckens 2008: 142)

- (3) & [Cyneheard] hine [= Cynewulf] *þær berad*
 and [Cyneheard] him [= Cynewulf] there *around-rode:indpst3sg*
 ond þone bur utan *beode*
 and the chamber from- *around-went*
 outside

‘And [Cyneheard] **surrounded** him (by riding) there and **surrounded** the chamber from the outside.’

(Petré and Cuyckens 2008: 144)

In predicative prefix constructions as exemplified in (2) prefixes are functionally equivalent to a secondary predicate: they are predicated of the theme of the clause, which in our example is the object of the transitive verb. Thus, the sentence can be paraphrased as “The army moved itself + it is IN.” In these cases, an isomorphism between the participant roles and argument roles exist which means that in our example a transitive verb remains a transitive verb when it is unified with the predicate prefix construction. Non-predicative prefix constructions are different, here the prefix cannot function as a secondary predicate. The authors note that especially prefixes whose etymon denotes route paths such as, for example, AROUND, OVER, ALONG express a prepositional relationship between the agent/subject/trajector and the location/object/landmark. Therefore, the sentences in (3) cannot be paraphrased by “He rode + the king was AROUND” and “He went + the chamber was AROUND.” Instead, the prefix *be*- and the root verb *ridan* and *gan* together express the (spatial) relation between subject (trajector) and object (landmark). The activity expressed by the verb and the path denoted by

the prefix leads to the direct object being totally affected by the complete and simultaneous surrounding. The authors point out that the notion of total affectedness is not part of the lexical semantics of the root verbs but of the construction itself which results in its non-predictability. In addition, a valency shift is entailed because when the intransitive verbs of motion *ridan* and *gan* are part of such a construction they become transitive and the construction's locative argument is linked with the syntactic function of the direct object. Petré and Cuyckens assume that non-predicative *be*-constructions are salient constructions, a property which clearly sets them apart from their prepositional cognates. The authors link the loss of the predicative *be*-construction (and other constructions of this type) to the OV/VO word order change where generally all kinds of verb specifiers like objects, adverbs, and prepositions occurring in front of the verb were replaced by new patterns following the verb. Since these constructions could be compensated by verb particle constructions without considerable loss of properties, they were gradually replaced. The non-predicative *be*-construction, however, could not be replaced by prepositional cognates because of their salient properties. More precisely, the specific properties of this construction, the affectedness component, and alternations in valency structure which also led to a number of further metonymic extensions (for details, see Petré and Cuyckens 2008: 150ff), preserved the construction during the time when drastic changes like the OV/VO change were taking place. The entrenchment of this construction made extensions to other roots, that is, nouns and adjectives possible, which led to formations like *besspectacled* (*spectacle*) and *befoul* (*foul*). Here, the prefix *be-* adds the sense "furnishing" to the construct without the support of a root verb. Interestingly, it is this extension that is most recent and most productive in PDE as the authors show by a quantitative study in the British National Corpus. They see this development as an instance of the grammaticalization of a construction in terms of their semantic and syntactic properties.

22.3.2 Case Study 2: From Compound to Suffix

In her monograph, Trips (2009) surveys in detail the rise of the suffixes *-hood*, *-dom*, and *-ship* in the history of English. She addresses the questions raised in the introduction which are repeated here: (1) Where does derivational morphology come from? (2) Are there general developmental stages that must be assumed for all derivational items? (3) What is the trigger for the development of derivational items? (4) Does the lexical semantics of derivational items determine the product of derivation?

In the following it will be shown that the three suffixes under investigation used to be nouns, that is, free morphemes, and developed into suffixes via a stage where they acted as morphological heads of compounds (relating to questions 1 and 2). Further, the properties of the nouns that developed into suffixes triggered the change from free morpheme to suffix (question 3), and their lexical-semantic development is part of the product of suffixation with these elements (question 4). Trips provides a lexical-semantic analysis based on Lieber (2004) for the insights gained on these aspects.

- These three nouns had different salient meanings: *hād* 'status, office, rank,' *dōm* 'authority, judgment' and *scipe* '(resultant) state, condition.' Functioning as heads of compounds and being relational nouns that required an argument, they predominantly occurred with nominal bases of the stage-level predicate type (cf. Kratzer 1995, Trips 2007, 2009). These properties are schematically represented in (5):

- Since they were all relational nouns, that is, two-place predicates, they required this relation to be fulfilled either in a syntactic phrase or a morphological structure. According to Trips exactly this property was a prerequisite for these elements to become suffixes since it implies the co-occurrence of the head of a phrase or morphological structure and its modifier.

(6) a. *-hood*[state] ‘a distinguishing feature of one’s personal nature’ (salient meanings “status, office, rank”; metonymies “state, time”)
 b. *-dom*[process] ‘possession of the qualities required to do something or to get something done’ (salient meanings “authority, judgment”; metonymies “territory, realm”)
 c. *-ship*[achievement] ‘result of a process of creating’ (salient meanings “created thing, (resultant) state”; metonymies “function, forms of address, skill/art”)

The polysemy of the suffixes partly arose by metonymic shifts: *-hood* predominantly occurred with nouns denoting persons which included the time dimension (e.g. *boy*, *child*). Therefore, the metonymic shift to “time, period” was possible. The salient meanings of *-dom* and its co-occurrence with nouns denoting persons led to metonymic shifts to “state, territory, realm” (if someone has authority, authority is generally restricted to a territory or realm). The salient meanings of *-ship* developed in the same way as those of *-hood*, and *-dom*: “state” can imply an office where duties are performed, and a person holding such an office may have certain skills and can be addressed by using a title. The three abstract features proposed by Trips—[state], [process], [achievement]—are based on an abstract dimension of aspect, and include all differences in meaning between the suffixes which result from their diachronic development.

The development of the three suffixes *-hood*, *-dom*, and *-ship* is analysed by Trips in Lieber’s (2004) lexical-semantic framework since it provides a systematic way to comprehensively describe the semantics of affixation. The main properties of the framework are: it is decompositional, cross-categorical, and it accounts for polysemy. Based on Jackendoff (1990), Rappaport Hovav and Levin (1992), and Mohanan and Mohanan (1999), Lieber assumes that a lexical-semantic representation consists of a Semantic/Grammatical Skeleton and a Semantic/Pragmatic body. The skeleton consists of a function and one or more arguments of that function, including the referential argument R (see further later) and contains a number of binary and privative features. Trips’s revised version adapted to diachronic aspects of affixation includes the following features: [–material] and [dynamic] refer to the core properties of *hād*, *dōm*, and *scipe*: since they are relational nouns denoting states or events intuitively involving doing something, they are also processual. Further, they denote bounded entities that can be composed of units ([+B], [+CI]). As introduced above, the features [state], [process], and [achievement] can be seen as a “diachronic imprint” since they include the salient meanings of the once free morphemes, represent the development of the suffixes, and define the differences between them. Moreover, they account for the underdetermination attributed to suffixes since they mark the semantic frame but not a specified concrete outcome. The lexical-semantic representations of the PDE suffixes *-hood*, *-dom*, and *-ship* are given in example (7):⁹

- (7) a. [–material, dynamic, +B, +CI, state ([R_i], [], <base>]
 hood
 b. [–material, dynamic, +B, +CI, process ([R_i], [], <base>]
 dom
 c. [–material, dynamic, +B, +CI, achievement ([R_i], [], <base>]
 ship

⁹ Contra Lieber, Trips assumes that the head is always the name of the function, i.e., the head is always the leftmost element.

The formation of complex words is subject to a mechanism of co-indexation. In the process of concatenating or juxtaposing two elements with two referents, the two referents are amalgamated into one referential argument by co-indexing these elements which share reference and interpretation. In this way, they are linked to a single constituent in the syntactic structure. Below the three developmental stages in the rise of the three suffixes are schematized (the several stages will be illustrated by *hād* but equally apply to *dōm* and *scipe*):

Stage I: *hād* = phrasal head

(8)	[−mat., dyn., +B, +CI, state ([R _i], [])]	[+ mat., −dyn., +B, +CI ([R _j])]
<i>hād</i>		<i>bisceopes</i>

At this stage, *hād* functions as phrasal head and requires an argument which is satisfied by the genitival modifier *bisceopes* in the syntax. The referential argument of *hād* “office” stands in a salient relation to the referential argument of *bisceopes* (in most cases it is the relation that modifies the office). Co-indexation is marked with different indices, i (head) and j (non-head). In the process of co-indexation the referential argument of the non-head is deleted which is indicated by underlining the non-head. At this stage in the development, co-indexation takes place at the phrasal level.

Stage II: *hād* = morphological head

(9)	[−mat., dyn., +B, +CI, state ([R _i], [])]	[+mat., −dyn., +B, +CI([R _j])]
<i>hād</i>		<i>bisceop</i>

During the second stage, *hād* is the morphological head of a compound. In line with Lieber, Trips assumes the mechanism of co-indexation to explain (complete) identification of reference of the two elements. The grammatical and semantic requirements are met since first of all, *bisceop* functions as an argument (office of bishop). Co-indexation of the highest head argument with the highest non-head argument takes place resulting in identification of reference. Second, the semantics of *hād* and *bisceop* are matched since an office is an abstract entity pertaining to persons and the argument matching its semantics is a noun denoting persons. In addition, both elements contain the feature [dynamic] since they both take part in the process of holding an office, and hence we gain an interpretation where the first element most suitably matches the semantics of the second element.

Stage III: *hād* = suffix

(10)	[−mat., dyn., +B, +CI, state ([R _i], [])]	[+mat., −dyn., +B, +CI([R _j], [])]
<i>hood</i>		<i>bishop</i>

In Lieber's framework, the difference between a compound and a suffix is hierarchical ordering, that is, the difference between juxtaposition and subordination. As can be seen from the representations given in (9) and (10), the only difference between the two structures is the hierarchical organization. This difference is also marked morphologically: juxtaposition is the concatenation of stem + stem, and subordination is the concatenation of a base and a bound element. We could also say that this manifests itself in the requirement of affixes to attach to a base which is part of the lexical entry of affixes. Therefore, there is a correlation of lexical-semantic representation and the morphological shape of the word.

Apart from the four questions posed in the introduction, Trips's diachronic account of the rise of suffixes also explains why the rivalry of suffixes does not exist: all meanings of PDE *-hood*, *-dom*, and *-ship* which developed in the course of time are part of the lexical-semantic representation of a suffix. Moreover, suffixes are sensitive to the meaning of their bases, regardless of whether bases denote a person and match the salient meanings of *-hood*, *-dom* and *-ship*, or whether they denote different properties and as a result match other meanings (mainly metonymies) of the suffixes. Under these assumptions, it is quite evident that parallel derivations are not synonymous. Although they form a paradigm building abstract nouns with a general meaning "state/condition of N" they also bear meanings different from each other, due to metonymic shifts. Thus, they are semantically different enough to coexist and cannot be called rivals.

22.3.3 Case Study 3: Language Contact as a Cause for the Rise of Affixes

In the introduction, different ways in which affixes arise in a language were discussed. Apart from internal causes, that is, the development of affixes out of free morphemes, external causes were addressed as well. Language contact as one such cause will be dealt with in the following.

Section 22.2 presented the history of the English affixal system, and it became evident that language contact especially with Romance languages like Latin and (Old) French impacted on English at different stages, predominantly in ME and EModE times. Many papers have commented on the language contact situation between English (ME) and (Old) French arising with the Norman Conquest of William I in 1066, and its linguistic consequences most drastically seen by the massive borrowing of loan words. Much less has been written on the structural impact this situation might have caused. In their paper, Trips and Stein (2008) define the borrowing of the OF suffix *-able* as an instance of structural borrowing since it is not "just" the integration of a new word-formation element that must be accounted for but also its syntactico-semantic and event structure properties that are part of its word-formation rule (their study is restricted to the deverbal pattern).

The main properties of *-able* formations discussed in the literature (cf. *OED*, *MED*, Marchand 1969, Dalton-Puffer 1996a, Delmas 2000) are: (i) denominal formations are

rare and not productive, (ii) deverbal formations are productive and have an active (“apt to V”) and passive (“apt to being V’ed”) meaning with the passive meaning being more dominant, (iii) *-able* formations are hybrid from the start which shows that it is analyzed and used as a suffix by English speakers/writers. These properties are corroborated by Trips and Stein’s study on ME and OF based on the *Penn-Helsinki Parsed Corpus of Middle English 2* (PPCME2) and the *Nouveau Corpus d’Amsterdam* (NCA). In ME, in the first sub-period (M1: 1150–1250) only one occurrence of a simplex with *-able* was found (*merciable*). In the following sub-periods, a constant increase of new (hybrid) types are evidence of the productive use of *-able*: in M2 (1250–1350) sixteen new types occur but they are still of OF origin. In M3 (1350–1420) twenty-one hapaxes are found, among them two hybrid forms (*knowable* ‘being known, able to know,’ *wilnable* ‘desirable’). Moreover, one text, the *Repressor of over-much weeting of the Clergie* by Reginald Pecock, with thirty-two hapaxes and eight hybrid forms (*forbedable* ‘that may be forbidden,’ *kutteable* ‘amenable to being cut,’ *makeable* ‘applicable,’ *markable* ‘remarkable,’ *unberable* ‘unendurable,’ *unsparable* ‘necessary,’ *wepeable* ‘regrettable’) clearly shows that the author productively uses the suffix to build new lexemes. In sub-period M4 (1420–1500) further new types and hybrids start to occur (e.g. *unspeable* ‘unspeakable’).

The process of adaption and assimilation of affixes was discussed in Section 22.2. If it is applied to *-able*, we may assume the following development: (i) First, a number of lexemes with *-able* were borrowed, these were analyzed as simplexes, for example *merciable* from M1; (ii) Although the element is recognizable as a foreign suffix, for example by a different stress pattern, new formations were built with native bases right away, implying that the element was analyzed as a suffix early on (e.g. *knowable* from M3); (iii) *-able* productively builds new formations with native verbal bases (of derivation) which is a sign of its full integration into the derivational system of English.

Concerning the semantics of formations with *-able*, in his French grammar from 1530 Palsgrave notes that the suffix can be combined with all verbs forming present participles in *-ant* and have an active or a passive meaning, like *mouvable* ‘apte or mete or able to chaunge, or to be chaunged’:

in our tonge we have none suche [word-formation element like *-able*], but must nedes use circumlocution by *apte*, *mete*, or *able*, and our infinityve mode; save that we have admitted as well adjectyves of the frenche tonge endyng in *able* and *ible*, as *commendable*, *visible*. (Palsgrave: 1530[1852]:303 in Trips and Stein 2008: 226)

In ME, on the model of OF, the meaning of *-able* can be active, that is, close to the present participle as in (11a), and passive as illustrated in (11b):

- (11) a. *This equinoxiall is clepid the gurdel of the first moeving, or ellis of the first moeivable. And note that the firste moevyng is clepid moevyng of the firste moeivable of the 8 speer, which moeving is from est to west, and eft ageyn into est.*

The final question briefly addressed is the systemic interplay of affixes. Put differently, which aspects in the native affixal system of English made it possible for *-able* to become rapidly productive in ME times? Two suffixes seem to have played a role: the native OE suffixes *-bære* 'productive of, carrying' and *-lice* 'being, characterized by.' In OE, the former builds denominal and deverbal adjectives from (*wæstm̥bære* 'fruitful,' *slæp̥bære*

‘soporific’). In other Germanic languages like Dutch (*-baar*) or German (*-bar*) this suffix builds deverbal adjectives in a highly productive way (e.g. German: *tragbar* ‘bearable,’ *ableitbar* ‘derivable’). Both languages also show a non-productive nominal pattern (e.g. German *fruchtbar* ‘fruitful’). In ME times, this suffix was lost and thus no longer available for deriving deverbal adjectives. The question is whether the loss of *-bære* can be attributed to French influence because a foreign suffix with similar meaning supplanted the native “equivalent.” The latter suffix, *-lice*, phonologically developed into ME *-lich/-li* and continued to build denominal (*freondliche*, *fleschly*) and deverbal adjectives, with active and passive meaning (*limpliche* ‘suitable,’ *untaleliche* ‘indescribable’). In the course of time, the deverbal pattern vanished. Dalton-Puffer (1996a) assumes an interaction between ME *-lich/-li* and *-able*: the English language was in need of a suffix building deverbal adjectives and *-able* was the right candidate to fulfill this need. By comparing English *-able* and German *-abel*, Lutz (2002) observes that in English the suffix allows many hybrid formations whereas German exhibits only non-native borrowings (e.g. *praktikabel* ‘applicable,’ *akzeptabel* ‘acceptable’). She also assumes that in Early ME formations with *-bære* were lost because they were replaced by formations with *-able*. According to her, this explains why in German the suffix *-bar* productively builds adjectives with a deverbal base and why formations with *-abel* are much less frequent and restricted to higher registers. Thus, although English and German used to have cognate elements to build deverbal adjectives, in English this possibility was lost due to the fact that OF impacted on ME to a high degree.

CHAPTER 23

DERIVATION IN A SOCIAL CONTEXT

LÍVIA KÖRTVÉLYESSY AND PAVOL ŠTEKAUER

23.1 INTRODUCTION

THE synergistic effect of linguistics and sociology has greatly contributed to the advancement of knowledge in various branches and subbranches of linguistics in recent decades. Therefore, it comes as a surprise that sociolinguistics has ignored morphology so systematically that there are only a few publications dealing with the social aspects of derivational morphology. Interestingly, while psycholinguistic research, especially research into the interpretation of novel compound words, has been massive in recent decades, sociolinguistics has very little to say about the social factors that affect either the formation or the interpretation of complex words. This chapter, therefore, maps the situation in this field and presents the available results.

Sociolinguistic aspects of derivational morphology may pertain to the following areas:

1. usage level;
2. system level.

Both of them may be studied both synchronically and diachronically.

23.2 USAGE LEVEL

23.2.1 Prieto

Let us start with a brief outline of research which concerns one narrow field of derivational morphology, in particular, that of evaluative morphology. Prieto

(2005: 3) confirms that unlike numerous sociolinguistic studies dealing with phonological and syntactic aspects of language there are just few publications addressing the issues of morphology, and that—in reference to Spanish—“there is a significant lack of formal references to the connection between some . . . suffixation processes and [their] sociolinguistic implications and effects on the interlocutors” (Prieto 2005: 10). This assessment of the situation in Spanish is of general validity. Some of the observations of Prieto and his forerunners in the field of evaluative morphology are interesting, but would need broader and cross-linguistic support. So, for example, Daltas (1987) and Giakoumaki (2000) observe that *females use more diminutives than males* which is accounted for by a more closer involvement of women with children. However, one may ask whether and to what degree the gradual recent changes in the Western family and the gradual acquisition of equal rights of women in the Western society influence the use of diminutives.

Prieto (2005) himself suggests (in reference to Spanish) that evaluative suffixes have had sociolinguistic effects, especially the effects of marking the language of social groups as well as the context (speech situation) marking. In relation to the former, he shows the differences between children, women, and different ethnicities/nationalities. In relation to the second type of effects, Prieto points out the relation between the use and/or non-use of diminutives and the *register*. Thus, for example, he shows that diminutives and augmentatives are primarily used in *informal contexts*, whereas superlatives may be both formal (mainly learned words) and informal. Prieto also discusses the inter-relatedness between the use of evaluative morphology and the social position of language users. Thus, for example, child/woman/low-class-related terms are more likely to be diminutivized than other terms. Prieto therefore concludes that “those segments of the society that traditionally have been bereft of influential power or status in the community at large are the ones that are more likely targets of diminutivization” (Prieto 2005: 186).

23.2.2 Säily

Säily (2011) studies the gender-related differences in the productivity of suffixes *-ness* and *-ity* in present-day British English. Based on analysis of the British National Corpus she concludes that “*-ity* is used less productively by women, while with *-ness* there is no gender difference” (2011: 119). Säily’s research gives support to the observation coming from gender studies that “[c]onsistent gender differences have emerged in synchronic variation . . . to the point that this has been called a sociolinguistic fact” (Säily 2011: 121).

Säily identifies the reasons for the different use of the *-ity* and *-ness* affixes in both diachronic and synchronic terms. Diachronically, the limited use of the *-ity* words was due to the “‘learned’ status and etymological foreignness of *-ity*” (Säily 2011: 122). By implication, men (with access to classical education) “preempted” the use of *-ity* words. Since these circumstances have been eliminated in modern times Säily seeks an explanation in the opposition between informational and involved

styles. The involved style is characterized, inter alia, by a high frequency of personal pronouns, and “-ness seems to be often used with pronouns...” All in all, the author demonstrates that what matters here is sociolinguistic variation in morphological productivity and that *gender* functions as a *robust social variable in English derivational morphology*.

23.2.3 Schmid

Schmid (2011), based on an analysis of the *Munich UCL Morphology Corpus*, is a more general study aimed at the evaluation of the use of new complex words in specific social situations by actual language users. Furthermore, Schmid studies the use of new words in various social groups, that is, the social contexts in which new words come into existence. Importantly, his sociolinguistic focus is on word-formation situations, individual language users as the actual coiners of new words, and the objectives of coining new words. Schmid points out, for example, that new words are coined because speakers cannot find the right word for the idea they want to express, or “because they want to express something in a particularly concise, pithy, original or humorous way. In press articles... the use of new words is often motivated by the need for brevity” (Schmid 2011: 73). Schmid illustrates this kind of nonce-formation with *50-minute* and *American-supplied*, and assumes that “these nonce-formations have more of a syntactic and co-textual function than a lexical, naming one,” because “both can be regarded as abbreviations and compressions of potential relative clauses (... *which lasts 50 minutes* and *which have been supplied by the US respectively*).” In addition, “the use of complex lexemes that are only institutionalized in parts of the speech community... signals group membership and identity when used by insiders, but can also function as an exclusion mechanism against outsiders [my emphasis]” (Schmid 2011: 102).

These three studies exemplify the immense diversity and the range of topics and methods available to sociolinguistic research into derivational morphology and, at the same time, indicate possible directions for future explorations.

23.3 SYSTEM-LEVEL

The system-level of sociolinguistic aspects of word formation primarily encompasses the following areas of research:

- (i) The influence of sociolinguistic factors like geographic circumstances, age, gender, education, profession, and language background upon the formation of new complex words. This is logical because each new complex word represents an individual naming act depending not only on the existing rules of the language

but also on a highly creative linguistic activity of the coiner reflecting his/her individual preferences which depend to a considerable degree on the factors mentioned above. This principle has been elaborated in Štekauer's conception of creativity within productivity constraints (Štekauer et al. 2005).

- (ii) The influence of sociolinguistic factors like geographic circumstances, age, gender, education, profession, and language background upon the interpretation of new complex words. The interpretation of novel words has been—as indicated above—extensively studied by psycholinguists, to a lesser degree by morphologists. What is missing are the publications examining the dependence of the interpretation of complex words upon the above-mentioned factors.

The sociolinguistic aspects of word formation, that is, the level of language system, concern the degree of employment, and hence, the productivity, of individual word-formation rules in forming new complex words. The social factors that can influence the formation/interpretation of novel complex words include age, gender, education, profession, and also language background in bilingual speakers. While research into the sociolinguistic factors affecting the interpretation of new coinages appears to be almost completely *terra incognita*, there are some fundamental studies by Štekauer et al. (2005), Keune et al. (2005), Keune et al. (2006), and Körtvélyessy (2009, 2010) that have established the foundations and, simultaneously, have indicated the directions of further research in this field.

Let us start with the concept of productivity. It has been studied extensively (e.g. Zimmer 1964, Aronoff 1976, Baayen and Lieber 1991, Aronoff and Anshen 1998, Baayen and Lieber 1991, a series of other articles by Baayen 1993, 1994a, Plag 1999, Bauer 2001, etc.) in recent decades. Štekauer (1998) points out that the Chomskian claim of productive syntax and low-productive word formation is unjustified if the two fields are put on the same cognitive foundations. Furthermore, the discussion of productivity has always been aptly connected to productivity constraints (e.g. Rainer 2005c). In addition to the “traditionally” adduced systematic language-internal constraints, including phonological, morphological, lexical, syntactic, and semantic ones, Plag (1999), Dressler and Ladányi (2000), and Bauer (2001) list some other, extra-linguistic factors, including:

- (a) Pragmatics (because of the denotation and connotation of some word-formation patterns these are not in common use, e.g. suffix *-some* in words like *twosome*. In principle, *twenty-fivesome* may be possible, but it is not usual because we do not usually operate with groups of 25 people).
- (b) Aesthetics, e.g. word-length.
- (c) “Accidents of cultural history,” e.g. while there is a noun *neglect* corresponding to the verb *neglect*, the noun *ignore* potentially convertible from the verb *ignore* is not commonly used.
- (d) Failure of hypostatisation: Coining a new word presupposes that there is an entity to be denoted by the new word. If there is no such entity, there is no need for a word.

- (e) Fashion, like the use of the suffix *-nik*. This factor is determined by the development of human society.

Interestingly, Dressler and Ladányi (2000: 108) include sociolinguistic factors in the general group of pragmatic factors, however, they do not explore them. It appears, then, that in spite of abundant literature on productivity constraints, there are at least two groups of factors that have been neglected and that deserve the attention of morphologists, in particular, the psycholinguistic and the sociolinguistic factors. In what follows a brief outline of the latter group of factors is provided.

23.3.1 Keune et al.

Papers by Keune, van Hout, and Baayen (2006 and ms.) appear to be (highly important) exceptions to the general absence of research into the sociolinguistic aspects of word formation. They examine socio-geographic variation in morphological productivity in spoken Dutch¹ by means of the Corpus of Spoken Dutch, totaling 8.9 million words. The method relies on Baayen's conception of productivity computation based on *hapax legomena*. The former research covers 72 affixes, the latter one 91 affixes. The sociolinguistic parameters include a geographical factor (Dutch in Flanders vs. in the Netherlands), education level (High vs. Non-High), gender (Women vs. Men), and age (Young: < 40; Mid: 41–60; Old: > 60). The data for all of them are available in the Corpus of Spoken Dutch. The research revealed significant effects for education level, sex, and age. The following is a brief summary of the results of this research:

- Affixes are used less productively by speakers aged between 19 and 40 than by speakers above 40. Highly educated older men (< 60) revealed the greatest overall productivity. This observation has been unambiguously confirmed by Keune et al. (ms.)
- While in the Netherlands speakers above 60 use affixes more productively, in Flanders they use them less productively compared to middle-aged speakers. This is explained by the authors by the fact that Dutch was not the official language in Flanders until 1963.
- In both the Netherlands and Flanders women use most affixes less productively than men (with the exception of diminutive suffixes).
- The geographical differences are also reflected in the productivity of affixes. The most productive affixes are the diminutive suffixes *-je* and *-ke*. While *-ke* is more productive in Flanders, *-je* is used more productively in the Netherlands.

¹ Keune et al. (ms.) also examine written Dutch (CONDIV corpus). The corpus does not, however, make it possible to reflect the influence of sociolinguistic data on word-formation productivity.

- The differences are not confined to the diminutive suffixes. The suffixes *-erig*, *-er*, and *-achtig* and the prefix *super-* are more productive in the Netherlands. On the other hand, the prefixes *her-*, *anti-*, *be-*, and *on-* and the suffix *-atie* are more productive in Flanders. Furthermore, the preference for certain affixes also appears to depend on age.

It should be noted that these two studies were preceded by a paper by Keune et al. (2005) in which the authors, *inter alia*, demonstrated the more frequent use of words derived by the suffix *-lijk* by speakers with a higher education level. Furthermore, the authors observed significant differences in the use of words in *-lijk* by men and women in the Netherlands and Flanders as a function of their education level. Keune et al. (ms.) examine affix productivity variation patterns on the basis of large spoken and written corpora.

23.3.2 Štekauer et al.

23.3.2.1 Theoretical Background

The next part of this chapter provides data obtained from extensive empirical research by Štekauer et al. (2005).

Štekauer et al. (2005) deal with the interrelation between the individual groups of language speakers and their preferred naming strategies in forming new words. They demonstrate that the naming strategy of each coiner may be influenced by a series of sociolinguistic factors, as indicated in the previous subsection. There are several central points in this respect which will be discussed in this section, including word formation as an individual act of naming, competition, productivity, constraints on productivity, creativity, and the opposition between two fundamental principles relevant (not only) to word formation, in particular, the principle of the explicitness (clarity) of expression and the principle of the economy of expression. All these factors are mutually interrelated and exert an influence on the strategies employed in any act of word formation. The basic results and correlations of empirical research will be summarized.

Strangely enough, the abundant literature on word formation and on the productivity of individual word-formation processes, rules, affixes, etc. is fairly abstract, giving an impression that the naming act is more or less automatic and that the language user is merely a passive user of language rather than its active creator. What has been widely ignored is the fact that each act of naming is a highly individual act. It depends on a language user's (coiner's) experiences, knowledge of the world, knowledge of their language, intellectual capacity, imagination, aesthetic values, education, age, professional interests, and language background. There are factors which can cause considerable variation in one's response to the naming task in terms of the selection of a naming strategy. In particular, the variation concerns the preference for more economical or more explicit linguistic representation of a concept of an extra-linguistic object to be named. Furthermore, it finds its reflection in more specific word-formation steps, like

the selection of a particular word-formation rule, a particular affix, etc. No doubt, each concept can be linguistically captured in different ways, by different strategies and rules. This situation engenders competition between word-formation strategies and inevitably endows the coining individual with a relatively high degree of freedom in coining new words. This freedom to choose a strategy, rule, etc., from a certain number of options is labelled creativity, a term which has been in use in derivational morphology in a different sense, meaning the extragrammatical formation of new words. We believe that the introduction of this term at the cost of polysemy is justified because it indicates a language user's conscious and deliberate naming activity enabling him/her to move within a space offered by the existence of competing naming strategies and productive word-formation rules. From this perspective, the notion of rivalry, more prevalent in the traditional, static views of word-formation, is just a part of a much more comprehensive concept of the act of naming, the concept whose focal point is the active role of language users. It is for this reason that the theory underlying the introduced research into the social aspects of word-formation is called creativity within productivity constraints. The aforementioned principles thus naturally establish the space to reflect the role of sociolinguistic factors in influencing the creative process of word formation.

From the perspective of this (briefly outlined) theoretical framework, each act of naming may be postulated to result from an intersection of the following factors:

- (i) The pressure of the productivity of individual Word-Formation Rules available for the naming of a particular concept.
- (ii) Various social factors like social stratum, education level, profession, age, gender, etc. It goes without saying that these factors affect the extent of actively and passively mastered vocabulary of individual languages users, and hence influence their linguistic behavior. This situation cannot—in our view—remain without effects upon the formation of new complex words.
- (iii) The extent of experience with a native language other than the mother tongue in bilingual speakers.

In accordance with Koch's (2001: 17) idea that the onomasiological viewpoint is closer to that of the speaker as a linguistic innovator than the semasiological viewpoint, Štekauer et al.'s research was implemented in the framework of an onomasiological theory (Štekauer 1998, 2005). Hence, the reader of this chapter is referred to the outline of its fundamental principles, especially in regard of the basic principles of the onomasiological theory and the theory of productivity in Štekauer (2005). For the sake of reader-friendliness, the following is a very brief summary of the most important principles relevant to this account. An onomasiological approach to word formation as proposed by Dokulil (1962) and further elaborated in Štekauer (1998, 2005) replaces the generally employed binary structure of complex words with a potentially ternary onomasiological structure which captures the motivating features and the crucial relations of the object to be named:

- (1) Determining constituent → Determined constituent → Onomasiological base
 of the onomasiological of the onomasiological
 mark mark

The onomasiological structure functions as a cognitive basis for the act of naming. At the onomatological level these three cognitively founded constituents are represented by morphemes of the particular language. This is called the *Morpheme-to-Seme-Assignment Principle* (MSAP). This principle makes it also possible to leave a particular cognitively founded member of the onomasiological structure formally unexpressed. The interrelation between the two levels establishes several naming strategies labeled in this approach as onomasiological (that is, naming) types:

Onomasiological type 1 (OT₁): all semes of the ternary onomasiological structure are assigned a morpheme in the naming act as in (2):

- (2) Result → ACTION → Agent
 novel write er

Onomasiological Type 2 (OT₂): the determining constituent of the onomasiological mark is left unexpressed:

- (3) Result → Action → Agent
 o write er

Onomasiological Type 3: the determined constituent of the onomasiological mark is left unexpressed:

- (4) Result → Action → Agent
 novel o ist

Onomasiological Type 4: the onomasiological mark is simple and unstructured, that is, it cannot be divided into the determining and the determined constituents.

- (5) Negation → Quality
 un happy

Onomasiological Type 5 (onomasiological recategorization) corresponds to what has been traditionally called conversion or zero-derivation, and the method of representation of semantic relations between the members of conversion pairs is illustrated by the following examples:

- (6) *switch*N—*switch*V: SUBSTANCE^{Instrument}ACTION
 Interpretation: Substance as an Instrument of Action

Given these five basic naming strategies available to a coiner, the influence of sociolinguistic factors may be examined in two ways. First, in terms of the strategy preferred for the coining of a new name for an object of extra-linguistic reality, and, second, in terms of the conflict between the economy of expression and the clarity (explicitness) of expression. The first tendency is listener/reader-friendly, because the meaning of a more explicit naming structure is more easily interpretable and predictable. The latter tendency favors the opposite universal feature of language, that is, the effort for the maximum possible economy of speech and is thus speaker-/writer-friendly.

From the point of view of the interrelation between the onomasiological and the onomatological levels, the economy vs. explicitness opposition results from the interaction of two types of opposition:

- (7) (i) Onomasiological level
 - (a) Onomasiological Types 1–3 (complex onomasiological structure)
 - (b) Onomasiological Type 4 (simplified onomasiological structure)
 - (c) Onomasiological Type 5 (absence of onomasiological structure)
- (ii) Onomatological level
 - (a) Onomasiological Type 1 (complex morphematic representation of complex onomasiological structure)
 - (b) Onomasiological Types 2 and 3 (economized morphematic representation of a complex onomasiological structure)
 - (c) Onomasiological Type 4 (economy due to onomasiological structure)
 - (d) Onomasiological Type 5 (absolute economy—no morphemic representation)

As indicated above, the present model distinguishes five Onomasiological Types ranging over all productive ways of forming new naming units. Since they are based on the criterion of which constituents of the onomasiological structure are linguistically expressed at the onomatological level, the determination of their respective productivities is an important indicator of the preferences of language users (or better, coiners) in terms of employing different cognitive processes underlying the act of naming, on the one hand, and the different ways of their linguistic representation, on the other.

Finally, it is postulated that the majority of naming acts can make use of a certain number of naming strategies as well as a certain number of word-formation rules that are productive, and therefore, permitted, within the scope of a particular cognitive category. Thus, for example, the cognitive category of Agent can be implemented by means of four general naming strategies:

- (8) OT1: *novel write-er*
- OT2: *teach-er, escape-ee*
- OT3: *novelist, librarian, milk-man, body-guard*
- OT5: *cheat_N*

All of these strategies can be employed for giving a name to a new Agentive entity. As it follows from the examples in (8), they range over the traditional, formally defined, word-formation processes of affixation, compounding, and conversion.² It is for this reason that we find our approach to the computation of productivity in word formation to be more objective than approaches that restrict themselves to the computation of the productivity of (competing) affixes. Our conception of competition is much broader because it encompasses all the available naming strategies within a single cognitive category.

By implication, the Productivity Rate is calculated as a proportion of individual naming strategies (onomasiological types) that produce complex words within a particular cognitive category (Agent, Instrument, Location, Causative, Diminutive, etc.). The share of the individual naming strategies of all the complex words within a particular category is expressed as a percentage.

Crucially, it is postulated that different social groups of language users classified by age, gender, profession, education, and language background, but also different psychologically classified groups of language users (e.g. inclinations for sciences vs. inclinations for humanities) tend to solve the particular naming task in different ways.

From the broader context of research into sociolinguistic aspects, the advantage of the method outlined is that it enables us to evaluate productivity as an implemented potentiality (capacity) reflecting the naming needs of a particular speech community.

23.3.2.2 *Experimental Research Results*

The experiment performed by Štekauer et al. (2005) included three naming tasks all of which concerned the cognitive category of Agent. The first set of tasks was based on making a selection from several options. Each object to be named was briefly characterized, for example, “a person who frequently interrupts other people when they are talking.” The characterization of the object of naming was followed by a set of options. In this particular case, they included *interrupter*, *interruptist*, *butt-in*, *butter-inner*, *cutter-in*, *cutman*, *interposer*, etc. As a second set of tasks, the respondents were asked to come up with their own words for an object to be named, for example. “Suppose that space aliens were about to land on Earth for the first time. What would you call a person who was supposed to meet them as a representative of the human race?” No naming options were offered in advance. The third naming task substituted verbal description with a drawing.

23.3.2.2.1 *Sample*

The sample of informants was divided into two groups, the native speakers whose parents were born in an English-speaking country, and native speakers whose parents were not born in an English-speaking country plus non-native speakers living in an English-speaking country. The former group (monolingual English speakers) were subdivided into various subgroups by occupation (students, educators, “other” professions), and by education (high school, college, graduate). The latter group (speakers connected

² And any other word-formation process used in the languages of the world.

with another language) was subdivided into groups based on the morphology of nouns in the particular language, namely, synthetic/agglutinative, synthetic/fusional, analytic/isolating, and polysynthetic. Due to very low numbers of informants (three), the polysynthetic group was not taken into consideration. The analysis was based on 1,300 responses from 145 respondents.

23.3.2.2.2 Results

The influence of sociolinguistic factors was evaluated in terms of the way the individual groups of respondents approached the contradictory tendencies in language, that is, the tendency to economy of expression and the tendency to explicitness of expression.

Influence of Occupation

The available sample of informants was eventually divided into three groups, in particular, students, teachers, and “other” professions. The sample included 60 students, 35 educators, and 50 “other” professions. The latter group of occupations was originally subdivided into those of civil servants, natural scientists and engineers, managers, manual workers, and medical doctors, but the data of all these sub-groups had to be cumulated into one set because of an insufficient number of questionnaires per sub-group. As a result, we obtained three basic occupational groups of comparable sample size, even though the last group was not a homogeneous one. In any case, by its structure, it significantly differed from the two other groups of language users, and thus its results were worth comparing.

The cumulative data covering all the tasks indicate that there are differences between the groups of students and teachers on one hand, and the “other” professions, on the other. They mainly concern Onomasiological Type 1 where the respective Productivity Rates are 53.86% and 57.30%, for the first two informant groups, and much lower in the “other” group (43.25%). This is, naturally, projected onto the situation in Onomasiological Types 2 and 3 where the Productivity Rates in the “other” professions are much higher. The tendency emerging from the data analysis is that while the language speakers belonging in the education-related area tend to form more comprehensive naming units, aimed at maximum explicitness and accuracy of the “labeling” of the objects of naming, the speakers belonging in the “other” professions prefer morphematically reduced ways of expression (economy of expression) (OT 2), and/or vaguer naming units with broader extension, the meaning of which is more difficult to predict (OT 3). All in all, there seems to be an obvious tendency indicating different strategies in the naming acts in the two groups of language users.

Education

In analyzing the questionnaires, our native speaker informants were divided into three groups, including those with high school, college, and graduate education. The data for the Onomasiological Type 1 show a rising curve in the direction towards higher education level. The Productivity Rate of Onomasiological Type 1 in the graduate group is noticeably higher than that of the high school informants. This is compensated for by the

higher Productivity Rate of the high school and college informants for the other three Onomasiological Types. The highest total Productivity Rate of Onomasiological Type 3 is attributable to the informants with the lowest education-level, with the Productivity Rate curve falling down towards the higher-educated speakers. The data indicate the preference of more highly educated language users for a more explicit way of expression, and the preference for a more general way of expression in the lower educated language users.

Language Background

The sample of informants encompassed 109 speakers living in an English speaking country, but born to parents coming from non-English speaking countries. A total of 109 informants provided us with 1,012 relevant responses for the category of Agent nouns.

The most noticeable difference between native speakers and non-native speakers concerns the respective roles played by Onomasiological Types 1 and 3 in these two groups of informants: *non-native speakers seem to lay greater emphasis on the explicitness of expression than native speakers*. It may be surmised that one of the reasons for this difference bears on the fact that since those informants who fall within the non-native group do not master English as fluently as native speakers their linguistic uncertainty makes them try very hard in the naming act to make their “products” as comprehensible as possible, and therefore, most explicit. Consequently, they prefer Onomasiological Type 1. Obviously, this is a possible psychological motivation behind this preference.

The second reason may be connected with the structural characteristics of the non-English languages that were shaping the linguistic behavior of the informants in the past. The hypothesis of the influence of language background seems to have been confirmed. English is predominantly an analytic language and therefore the results obtained from native speakers should be closest to those obtained from the isolating/analytic group of background languages (seventy-three informants). The data appear to confirm this position, though the differences do not rise to statistical significance. There is a striking similarity between the naming tendencies in this group of informants, while the agglutinative (nineteen informants) and the fusional (seventeen informants) background languages deviate from the “native” data in a noticeable way, as reflected in their lower p-values in the chi-square test. This primarily concerns the role played by Onomasiological Types 1 and 3. Furthermore, it is of no surprise that the Productivity Rate of Onomasiological Type 1 in the agglutinative group is the highest of all. This may be explained—in addition to the psychological reasons relevant to all three groups of non-native language background—by the morphological characteristics of agglutinative languages, aiming at expressing complex morphological meanings within one word.

23.3.2.2.3 *Conclusions*

In general, the naming strategies of the two basic groups of speakers seem to differ because *non-native speakers seem to lay even greater emphasis on the explicitness of expression than native speakers*. The influence of *linguistic background plays its role in*

the naming strategies of non-native speakers. Although the limited sample of informants with “Germanic linguistic background” does not enable us to draw any indisputable conclusions, the agreement of the results between them and the native English speakers in terms of almost identical preference for affixal types is remarkable.

It has also been shown that any assessment of the influence of any of the above-discussed sociolinguistic factors should be related to the specific *cognitively founded category* (Agent, Patient, Instrument, etc.).

Finally, the informants’ assessment of the probability of occurrence of the selected options provided us with a relatively high number of the “likely” responses. This acknowledges their feeling for a *creative approach to naming*. This *mainly applies to the group of students* who most readily accept unconventional naming units that do not correspond to productive rules of word formation. This does not seem to be a surprise, and might be accounted for psychologically by the dynamism of the young generation compared to the more conservative generation of their parents. Moreover, the share of “extremely unlikely” answers in the university-educated informants is much higher than in the lower-educated groups, which suggests that *the awareness of grammaticality in higher-educated speakers is stronger*. The tolerance to “creativity” (even the creativity that trespasses grammaticality) characteristic of the young generation is also typical of speakers with lower education, even if the reasons underlying this fact may partly differ in these two groups of speakers.

The native vs. non-native comparison shows that the pressure of productive Word-Formation Rules is perceived by native speakers to be a little stronger.

All in all, the results of Štekauer et al. (2005) bear out the hypothesis of *sociolinguistic conditioning of the individual acts of word-formation*.

23.3.3 Körtvélyessy (2009, 2010)

In her extensive research project, Körtvélyessy (2009, 2010) relied on an onomasiological approach and a similarly structured questionnaire aimed at two target groups of bilingual speakers: Hungarians living in Slovakia (142 respondents, 1,195 responses) and Hungarians living in English-speaking countries (146 respondents, 1,252 responses). These groups of informants were selected in order to examine the influence of Hungarian as an agglutinating language upon dominantly inflectional Slovak, on one hand, and upon dominantly analytic English, on the other. The data for both of these groups was also evaluated in terms of age (<18, 19–24, 24–39, 40–59, <60), gender, occupation (engineers, IT-workers, sciences; lawyers, journalists, teachers, administration; manual workers and artists; managers, businessmen; housewives, pensioners, the unemployed), and education (primary school, high school, college, university).

23.3.3.1 Results

The non-canonical correlation was used for data analysis because it enables us to search for relations between more than two sets of variables.

23.3.3.1.1 Gender

Neither English nor Slovak data show significant influence of gender on the choice of the onomasiological type. Since this result has been confirmed for both groups of informants it can be assumed that the influence of gender on productivity in word formation is not significant, especially in comparison to the other factors like age, education, and occupation.³

23.3.3.1.2 Age

The influence of age is evident. The lower the age of the Hungarian-English informants the stronger the tendency towards complex words with simple onomasiological structure or non-transparent complex words (e.g. *billbold* as a name for a person who installs billboards using a scaffold (informant aged 10) vs. *billboard placer* (informant aged 45); a person addicted to phoning: *phoneaholic* (informant aged 9) vs. *phone addict* (informant aged 39). At the same time, the Slovak-Hungarian informants of a younger age prefer borrowings that are not transparent in the Slovak language (*billeš*,⁴ informant aged 12). The influence of age can also be observed at the level of word-formation type⁵—the younger age categories prefer word-formation types that are not typical of the given conceptual category (e.g. *bordnoha*,⁶ Quality-Patient); *scaffold* Instrument-/Action/-Agent. It can be explained by their effort aimed at unconventional, creative solutions. This phenomenon has been observed in both groups of informants and it can be generalized as a phenomenon typical of the relationship between word-formation type and age.

23.3.3.1.3 Education

The non-linear canonical correlation shows a strong association between the factors of age and education. The higher the educational level of the informants the stronger the

³ This observation is partly called into question by Kečkéšová (2012) who tested two target groups—55 high school students and 45 kindergarten children for their preferences within a sound symbolism task. Inspired by Berko (1958), Kečkéšová's questionnaire included potential words, such as *mip* and *map*, and the respondents were asked to match these words with pictures that were supposed to represent the meaning of the individual potential words according to their sound symbolic properties, including small vs. big, light vs. dark, angular vs. round, fast vs. slow, narrow vs. wide. The former members of these pairs were supposed to represent a diminutive meaning (symbolized in the potential words by the high front vowel), the latter members an augmentative meaning (represented in the potential words by a back vowel). The results indicate that while high school female users of language are more sensitive to sound symbolism than their male counterparts (68% vs. 45%) there is hardly any difference between the kindergarten children (65% vs. 67%), which might indicate the correlation between the parameters of gender and age. Certainly, these results will need support from further experimental projects of a similar sort (sound symbolism) as well as from additional experiments concerning word formation in general.

⁴ Motivated by a combination of *bill* + Hungarian agentive suffix *-es*.

⁵ The Word-Formation Type in this approach is determined by the onomasiological structure underlying a complex word.

⁶ I.e., board + leg.

tendency towards a more transparent onomasiological structure of the coined complex words.

23.3.3.1.4 *Occupation*

The influence of occupation correlates highly with the parameters of age and education. From the perspective of occupation, the most creative group in the field of word formation seems to be the group of students. Both Slovak and English students' word-formation strategies tend to deviate from the established patterns. Similar deviations can be observed in the category of manual workers and artists.

23.3.3.1.5 *Language Background*

The non-linear canonical correlation has confirmed the fact that language background is a completely independent sociolinguistic factor which differs from the rest of the sociolinguistic factors examined. While strong correlations are found between the factors of age and education, education and occupation, and occupation and age, no similar correlations are observed for language background.

The research has confirmed the influence of language background on naming strategies. Lower p-values in the canonical correlation proves the statistical significance of the experiment in question. While the productivity of onomasiological types for Slovak and English differs, Onomasiological Types 1 and 3 are the most productive onomasiological types in both languages. It is beyond dispute that the Hungarian language shared by both groups of respondents may be held responsible for this similarity. The canonical correlation has confirmed this observation, too. In addition, in both groups of informants the following tendency can be observed: the stronger the influence of the Hungarian language⁷ the higher the productivity of Onomasiological Type 2, which indicates that the Hungarian language does not favor a detailed motivation in the naming strategies; instead it appears to prefer the brevity of expression.

Similar results were arrived at by Klembárová (2012) who compared the word-formation strategies of monolingual (Slovak) and bilingual (Slovak-Hungarian) informants in a series of experiments (matching, naming of a picture, etc.). The experiment included twenty-six bilingual and twenty-three monolingual 10-year-old children and twenty-four bilingual and twenty-three monolingual adult respondents. The results suggest that Onomasiological Type 3 (correlating with suffixation) is the most favorable strategy in both the monolingual and the bilingual groups. The difference has been observed in the second most frequent onomasiological type: while the bilingual

⁷ The strength of influence of Hungarian was evaluated by means of the language background information provided by the informants about themselves and their parent's language (a five-degree scale ranging from "I am a native speaker" down to "I know a little bit, but I have a hard time to communicate normally in the language") and about the language(s) spoken at home (a five-degree scale ranging from "always" to "never"). Based on these data, the informants were divided into language background groups, e.g. "Slovak+Hungarian" means that—given the specific language background—the influence of the Slovak language is stronger.

respondents prefer Onomasiological Type 2, the Slovak monolingual respondents prefer Onomasiological Type 1. In general, bilingual respondents form more explicit potential words.

23.4 INTERPRETATION OF COMPLEX WORDS

In reference to various incorrect interpretations of compounds, such as *house-bird glass*, Gleitman and Gleitman (1970) relate their misinterpretation to the *educational level* of language users. Their informants fell within three different educational groups: (a) graduate students and PhDs in various fields; (b) undergraduates and college graduates; and (c) secretaries with high school degrees. In many cases, their informants from the group of secretaries proposed various “unacceptable” readings which corresponded to the compound *glass house-bird*, *glass bird-house*, or a paraphrase like *a house-bird made of glass* (in contrast to PhD informants who avoided such errors). The analysis of their research results made Gleitman and Gleitman conclude that there were “very large and consistent differences among these subjects of differing educational background” (1970: 117) and that “[t]he less educated groups make more errors, and to a significant extent make different errors than the most-educated group” (1970: 128). Thus, Gleitman and Gleitman assume that “giving language judgments—retrieving and making use of one’s intuitions—is relatively hard, compared to talking and understanding” and is accompanied by “extensive individual and population differences in performance of harder judgmental tasks...” (1979: 107). While the research of the Gleitmans concerns the predictability of meaning, that is, the interpreter’s perspective rather than the coiner’s perspective, it indicates that the level of education may play a role in word-formation, in general, and in the productivity of word formation, in particular. No wonder, the productivity of Word-Formation Rules appears to be one of the factors influencing the predictability of novel complex words (cf. Štekauer 2005).

23.5 CONCLUSIONS

Very little has been done to examine the influence of sociolinguistic factors on word formation. Therefore, let us conclude this chapter with programmatic words by Ryder (1994) in the chapter of her book entitled ‘Areas for Future Research’. Notably, she stresses the significance of extensive research in the field discussed in this chapter. Ryder maintains that

it is time to realize that linguistic variation is not confined to a few pronunciation and lexical differences, and to begin investigating what its parameters really are. As a start, I would like to look at compound interpretation in groups that vary

systematically in age, sex, educational background, and possibly profession to see if there are predictable correlates between compound interpretation styles and any of these other factors. (Ryder 1994: 199)

These words apply not only to the area of complex word interpretation but also to the area of word formation.

CHAPTER 24

ACQUISITION OF DERIVATIONAL MORPHOLOGY

EVE V. CLARK

As children begin to master a vocabulary, they typically attempt both simple and derived word-forms from very early on. All early words in production are unanalyzed, so whether they are simple or derived makes little difference. Indeed, among words children pick up early in English are many denominal verbs, with no affixation (e.g. *to comb*, *to bicycle*, *to fish*, *to picnic*). They also pick up a variety of other complex word-forms in their first years of talking, from *breakfast*, *somersault*, and *tallest*, to *mountaineer*, *teacher*, and *complicated*. In order to study children's (and adults') knowledge about derivation, one needs to look at (a) spontaneous analyses of word parts and elicited glosses that indicate understanding of novel word meanings, and at (b) novel coinages in their spontaneous and elicited production of language. Simply tracking and counting all the conventional words children say that contain an affix of some kind does not tell one directly what children know about derivation per se. Children pick up words wholesale as they acquire a vocabulary, so they often make use, early on, of many an established word that contains an affix of some kind, some productive and some not. Compare the affixes in *washing-machine*, *skier*, and *bicyclist*, with those in *strength*, *kingdom*, and *motherhood*. Mere production of established words does not demonstrate knowledge of the parts, here the affixes *-ing*, *-er*, and *-ist* compared to *-th*, *-dom*, and *-hood* (where the latter three are less productive; but see Lieber 2009b). For children, all these words start out as unanalyzed chunks, and are only later unpacked or analyzed into parts. The parts—roots and suffixes, here—must also become linked to a specific meaning for each element.

Knowledge of a derivational affix is impossible to assess from observations of established words alone. Children may recognize fairly early on that some words they use form small families with such near neighbors as in *washing-machine*, *wash*, and *washing*, or *skier*, *skis*, *skiing*, and *skied*, or *bicycle*, *bicycling*, *bicycled*. Eventually, they should also

recognize the *-er* in *skier* as being the same as the *-er* in *swimmer*, *teacher*, and *writer*. How soon do children start to gather such groups of related words together? How soon do they recognize relations of form and meaning among the words they are acquiring? Can this kind of knowledge be readily tapped? The problem is evident when we consider another small family—*strength*, *length*, *breadth*, and *health*. The suffix *-th* is no longer productive, and the adjectival forms it was attached to, to construct these nominals, are pretty distant from their adjectival sources—*strong*, *long*, *broad*, and *hale*. At what age can speakers of English identify *-th* as a possible suffix? How many instances of nominals in *-th* do they need to hear? And even then, do they necessarily identify the final segment as a nominal suffix? They may have an easier time, though, with affixes that are currently productive in the language—affixes like English *-ie* (diminutive), *-er* (agentive, instrumental), both widely used, or *-ness* (nominal), *-ful* (adjectival), or *-ly* (adverbial), also all well represented in the established lexicon.

In this chapter, I will review children's knowledge of derivation in word formation, what different studies can tell us about order of acquisition within a language, factors that influence that order of acquisition, and consistent differences in what children understand compared to what they can say.

24.1 ACQUIRING DERIVED FORMS

What do children know about the morphemes that contribute to word meanings as well as forms? One source of information is what children have noticed about the words they hear or produce, typically in the form of spontaneous analyses of words into parts, for example, *A horsie is a little horse*. Such analyses show that children can distinguish the root or core word stem from the affix (*horse+ie*), but their spontaneous analyses are not particularly frequent although they do indeed occur (see, e.g., Clark 1978, Slobin 1978). They also depend on the child's knowing enough language to be able to offer a gloss.

Another approach to assessing what they know has been to present children with unfamiliar, novel words, and then ask them what these might mean. If the children can analyze the word-forms they hear into root and affix, say, and if they understand the meaning of the affix, they can often offer a gloss or interpretation. For example, asked what a boy who's a *jumper* could be, one 3-year-old readily replied "He jumps things." Notice that to arrive at this interpretation, the child must remove the affix *-er*, turn the root back into a verb with an added 3Psg inflection, and then supply it with a singular subject, *he*, and an indeterminate direct object, *things* (Clark and Hecht 1982). This suggests that this 3-year-old understands that *-er* marks the agent of an action of jumping. If children can identify affixes as adding specific meanings to root word-forms, this provides clear evidence that they are able to analyze the forms in question. Then, in addition, if the same children can make use of the affix *-er* to construct innovative verb-root+affix combinations themselves, to express new meanings, we have further evidence that they have acquired the derivational suffix *-er* as an agentive suffix

(see Clark and Hecht 1982, Mulford 1983, Clark and Berman 1984, Clark 1993). Once children can do this, these affixes are productive for them, and their meanings can contribute to a growing family of words that share that affix as part of their meaning. So in English, the word-shape Verb+*er* offers a template for constructing nouns from verbs to mean “the agent of the action denoted by the verb.”

In short, both the ability to show understanding of an affix in unfamiliar words and the ability to produce new word-forms with that affix added to other verb-roots provide strong evidence that children have acquired the relevant knowledge about deriving new agentive nouns in English. And this provides stronger evidence for knowledge acquired than merely saying or repeating an established word that happens to contain a particular affix.

24.2 ORDER OF ACQUISITION

English-speaking children appear to pick up several productive affixes very early. The first to emerge in their speech, well-modeled by parents and other adults, is the diminutive *-ie*, added to monosyllabic nouns, as in *horse/horsie*, *bear/bearie*, *dog/doggie*. This is supported perhaps by the */-i/* suffix on such established forms as *kitty* and *puppy*, as well as the address terms *mommy/mummy* and *daddy* (e.g. Ferguson 1964). Diminutive forms in other languages also tend to be acquired early by young children, probably because of extensive uses of diminutive forms in parental speech (see, e.g., for Dutch: Gillis 1997; Finnish: Laalo 1998; Greek: Stephany 1997; Italian: Bates and Rankin 1979; Quiché Mayan: Pye 1986; Polish: Haman 2003, Dabrowska 2006; Russian: Kempe et al. 2003, Kempe and Brooks 2001; Spanish: Melzi and King 2003, King and Melzi 2004; see also Dressler 1994). However, very young children may not use diminutives productively: rather, they *acquire* diminutive word-forms, and only later come to analyze them and use the diminutive affix productively (see Svaib 1993).

The English adjectival suffixes in *-y* and *-ed* also emerge early on, as 2-year-olds use innovative forms to describe both inherent properties (e.g. *crowdy* [= crowded], *nighty* [= dark]) and temporary states (e.g. *crumbed* [= covered in crumbs], *floured* [= covered in flour], *fastered up* [= speeded up]) (Clark 2001). One sees a similar use of past participial adjectives for resultant states in French children's early uses of *c'est+PP* (Veneziano and Clark 2012).

Next in line to emerge in English is the agentive (and instrumental) suffix *-er*. This appears with regularity first only in agentive nouns like *cooker* (for conventional *cook*) or *lifter* (for “someone who lifts things”), and then in later-emerging instrumental nouns as well, such as the innovative *climber* (for *ladder*) or *hanger* (for *hook*) (Clark and Hecht 1982, Clark 1993). The next derivational morpheme to emerge is the negative prefix *un-* on verbs, used to convey the reversal or undoing of an action (*lock/unlock*) or the meaning of privation (*saddle/unsaddle*).¹ However, children take some time before

¹ This prefix also appears on adjectives to express the opposite semantic value, as in *happy* vs. *unhappy*.

they make use of negative *un-* in their own innovations (Clark et al. 1995). By age 4 to 5, English-speaking children begin to master another agentive suffix, *-ist*, used in such child forms as *trumpetist* (for *trumpeter*, in parallel to such forms as *pianist*, *flutist*, and *violinist* (see Clark and Cohen 1984).

Most other derivational affixes in English are not acquired until after children enter elementary school, and some are only mastered once children have access to another source of information besides speech about the relations between such pairs as *danger*–*dangerous*, *library*–*librarian*, *magic*–*magician*, or *volcano*–*volcanic* (e.g. Deacon and Bryant 2005, 2006). Stress-shifts with their concomitant changes in the value of the affected vowels, as well as consonantal shifts from /-k/ to /-ʃ/ make the relations between base words and such derived forms as adjectives in *-ic*, *-al*, or *-ous*, or nouns in *-ation* or *-ity*, for example, hard for children to discern (see, e.g., Lewis and Windsor 1996, Jarmulowicz et al. 2008).

What children acquire when also depends on which language they are learning. In some languages, derivation is king for the construction of new words. In others, derivation shares that role with compounding, and in others still, compounding provides the major option for lexical innovation. What is favored in any particular language depends largely on typology. For example, in Germanic languages, we see both compounding and derivation used in the construction of new words, with compounding somewhat favored over derivation, but in Romance and Semitic languages, derivation is heavily favored for most lexical innovations. One result of this is that the productive options displayed to young children in the language they hear from adults will differ from one language to the next, and this affects the kinds of lexical innovations observable in children.

Finally, several factors appear to affect what children learn first, and hence their general course of acquisition for constructing new words. One that has already been mentioned is *productivity*, here the productivity of the affixes available. In general, more productive affixes are more frequent—they occur on more stems (types) and hence appear on more tokens in everyday speech. This should make them more available for children to identify and analyze in the adult speech addressed to them. But how to measure productivity remains an issue. One approach has been to look at the frequency of different affixes in the established lexicon (e.g. Baayen and Lieber 1991), but while this offers a good record of past productivity extrapolated to current usage, it may not measure the particular word-forms that adult speakers currently favor in their innovations. Another approach has been to look at normative recommendations for which word-formation pattern(s) to use for specific meanings, but speakers often depart from such recommendations (e.g. Berman 1987). And another has been to focus on those word-forms adult speakers actually use in their lexical innovations, spontaneous and elicited, viewed as *natural productivity* (see Clark and Berman 1984; see also Schultink 1961).

Two further factors are *semantic transparency* and *formal simplicity*. Before children can make use of any derivational affixes in their own novel word formations, they need to have identified an affix as an affix, and analyzed its meaning in the forms

where it appears. Once they have done this, children can add that affix to their repertoire of devices available for use in the construction of new words with the relevant meaning-type. In other words, the affix has become *transparent* for the child in the meaning it adds to a stem, and hence available for constructing a new word-form to convey a new meaning. *Simplicity* of form in new words is also a factor early on. If children can construct a new word with minimal, or even no, changes in the form of the source word, they do so. This seems to make some word-formation devices highly accessible very early, with the result that children learning languages with conversion make early use of that option in their lexical innovations.

24.3 USING DERIVATIONAL AFFIXES IN LEXICAL INNOVATIONS

As children learn words, they gradually build up a vocabulary. At age 2, they have typically mastered between 200 and 600 words for production (a good number more in comprehension). Their repertoire gets steadily bigger as they get older, and by age 6, their comprehension vocabulary is typically estimated to be around 14,000 words, with somewhat fewer in production (Clark 2009). But children often try to talk about things for which they have yet to acquire the conventional terms, so they coin new words on such occasions that later yield to established adult forms. These coinages, both spontaneous and elicited, offer invaluable information about what children know about the internal structure of words and about the meanings of the component roots and affixes.

24.3.1 Conversion

In some languages, conversion or derivation without any affixation added (sometimes called zero derivation) provides a ready means for constructing new verbs from nouns, as in English. Young children make early use of this option in some languages, possibly because the change of word class without any affixation is so simple. Such verbs, for adults, fall into a small number of categories; within each category, they lie on a continuum from well-established verbs at one end to innovations at the other; and to use denominal verbs, speakers of English rely on a convention that places constraints on the actual process of innovation (Clark and Clark 1979). Children acquiring English readily form new verbs from nouns, as shown in Table 24.1, as do children acquiring other West Germanic languages like German, Dutch, and Swedish (Clark 1982, 1993). In all four languages, children as young as age 2 favor, as the source nouns for new verbs, words for instruments, objects (*locata*), and locations—common adult categories of denominal verbs in those languages.

Table 24.1 Some innovative denominal verbs*Instrument verbs*

- 1 S (2;4, wanting to have something weighed): *you have to scale it first.* [= weigh]
- 2 S (2;7, having hit his baby sister, explaining why she was crying): *I broomed her.* [= hit with a broom]
- 3 D (2;9, of a sock): *did you needle this?* [= mend]
- 4 D (2;9.10, after mention of boats): *and we might see a man oaring a boat with oars.* [= row]
- 5 S (2;11, not wanting his room to be cleaned): *don't broom my mess.* [= sweep]
- 6 D (3;2.9, picking up a Cuisinart blade from the sink)
Mother: You shouldn't take that. It's very sharp.
D: *but I didn't blade myself.* [= cut]

Locatum verbs

- 7 DH (2;3, talking about getting dressed): *Mummy trousers me.* [= puts on my trousers]
- 8 J (2;6, asking a teacher to toss a pillow at him in a mock pillow-fight): *pillow me!* [= throw a pillow]
- 9 EB (3;4, deciding not to wear her new nightgown outside in the patio): *mine will dust.*
Mother: Dust?
EB: *mine will get dust on it.*
- 10 CB (3;11): *I'm crackering my soup.* [= putting crackers into]

Location verbs

- 11 CA (4;6, waving a funnel at her younger sister): *I'm going to funnel you.*
Ff, ff, ff. You're all in there.
- 12 SO (~5;0): *I'm going to basket those apples.* [= put into a basket]
- 13 CB (5;5, asking Mo to stop her sister from putting popcorn "beads" on her thread): *Mom, will you keep Eva from threading on mine?*
- 14 K (7;0, asking Mo to wrap a towel round her as she got out of the bath): *towel me, mommy.*

In other languages, children make much less use of the conversion option and, if they use it, do so only at a later age. For example, children acquiring French do construct some new denominal verbs, always placing them in class 1 (verbs in *-er*, the class that accounts for around 90% of all French verbs), but this is relatively rare before age 4. Some typical examples include innovative *pincer* 'to paint,' from *une pince* 'paintbrush' (used in place of adult *peindre*); *mètrer* 'to measure,' from *un mètre* 'ruler, tape-measure' (used in place of adult *mesurer*), or *pantoufler* 'to put on slippers,' from *une pantoufle* 'slipper' (see Clark 1982).

Children acquiring Romance languages make little use of conversion to form new verbs, but they do rely on it rather earlier in forming new nouns (from verbs), as in *le tomber* 'the fall/ event of falling,' from *tomber* 'to fall,' or *le jeter* 'the throw/ event of throwing,' from *jeter* 'to throw.' Most nouns from verbs like these

denote activities or events, and are not common among children's early coinages (Clark 1993). Although Hebrew has historically relied on present-tense (*benoni*) verbs as a source of nouns, as in /ʃofet/ 'a judge; judges-3Psg' or /mocec/ 'a pacifier; sucks-3Psg,' young children make relatively little use of such conversion in their early coinages (Berman 1985). In Hungarian, though, they do construct some early nouns from verbs and adjectives, as in *savanyú* 'a grape,' from the adjective "sour" (for adult *szôlo*), or *farag* 'knife,' from the verb "to carve" (for adult *kés*) (MacWhinney 1978).

Although conversion would appear to offer children a very simple way to construct a new verb or a new noun from an existing word, its availability to young children depends on how productive this option is in the language. Moreover, as an option for new-word formation, it is not equally available across languages. Some languages make extensive use of it, as English does—for constructing new verbs—but others seldom use it, instead favoring derivational affixes.

24.3.2 Derivational Affixes

Before children can start to make use of derivational affixes in constructing new words with the relevant meaning added by the affix, they must learn to identify affixes apart from the stems they appear with in established words, and analyze their meanings. How soon children do this depends on the frequency of the established words containing each affix, and on how productive that affix is in adult speech. Some derivational suffixes emerge in children's production early on, typically beginning between age 2 and 3, while others are still not mastered at age 10.

Overall, languages are biased towards suffixes (Cutler et al. 1985, Hupp et al. 2009). Most inflections appear as suffixes, as do most derivational affixes. This probably reflects a processing preference or bias, whereby the core word or stem is readily recognized from its onset, especially when the onset is not obscured by any prefixed material (see Kuczaj 1979, Cole et al. 1989, Lahiri and Marslen-Wilson 1991). Once a core stem has been identified, it should then be easier to process any modulations of meaning contributed by derivational suffixes added onto the end of the core stem. The presence of prefixes, however, may well delay children's recognition of the pertinent core word or stem, and so make it harder for them to identify both the core forms and specific prefixes (see, e.g., Pye 1986, Mithun 1989).

24.3.2.1 Diminutive Suffixes

Diminutive suffixes appear early, but mainly in conventional diminutives such as English *kitty*, *dollie*, and *birdie*, as well as terms like *doggie*, *horsie*, and *bearie*. Children sometimes notice the diminutive ending and extend it, as when Leopold (1949) observed his daughter, aged 3;3, adding the English diminutive *-ie* to a variety of nouns not normally so modified, for example, *wallie* (wall), *chairie* (chair), *lappie* (lap), and

even the plural *books-ies* (vol. 4: 45). In my diary data, I observed similar uses, as in the following exchange:

- (1) D (3;6.25, out of the blue one morning): *These are my pantsies. D'you know what I call a cat?*
 Fa: What?
 D: A cattie. *D'you know what I call a bear? A bearie!*
D'you know what I call a dog? A doggie!

But when asked what they could call a “very tiny wug,” Berko (1958) found that few children aged 4 to 7 could supply a diminutive ending. In a more extensive elicitation task, Svaib (1993) found that 4-year-olds produced significantly fewer diminutives in *-ie* than 5-year-olds or adults (24% vs. 48%, 49%). When all the diminutive affixes used were counted, 4-year-olds still produced fewer (32%) than 5-year-olds (62%) or adults (83%). An analysis of diminutive usage in four children in CHILDES showed that they used very few diminutives as they went from age 2 to age 4, and they used them on only a very small number of lexical items. But when Svaib asked 4-year-olds to pretend they were talking to a 2-year-old as they provided labels for pictures, they produced many more diminutive forms (49%, up from 24%). Diminutive use may therefore be largely register-dependent in English, with 4-year-olds able to form diminutives in the appropriate context.

Children acquiring Italian show mastery of several diminutive suffixes by age 2;0, but they use them only for small size, and take considerably longer to master other uses to connote value, affection, humor, and sarcasm in Italian (Bates and Rankin 1979, King and Melzi 2004). Diminutive affixes are also very frequent in languages like Russian, and some researchers have argued that their use in adult speech to children also aids in the acquisition of noun gender (Bogoyavlenskiy 1973, Kempe and Brooks 2001).

24.3.2.2 Agentive and Instrumental Suffixes

Suffixes for agentive meanings start to emerge in English in children's innovations around age 3, and become moderately well established by age 5. However, their acquisition continues well into the school years as children learn to read and make use of morphological information to interpret unfamiliar words (e.g. Sterling 1982, Wysocki and Jenkins 1987, Tyler and Nagy 1989, Windsor 1994, Lewis and Windsor 1996, Deacon and Bryant 2005, 2006, Roy and Labelle 2007, Rabin and Deacon 2008). The use of affixes to express agentive meaning in words newly constructed lags behind children's earlier, extensive reliance on compounding for this purpose (Berman 2009). Indeed, in English-speaking children, compound nouns typically account for 70–80% of their innovative nominals up to age 4 (Clark 1993, Becker 1994).

While agent nouns in *-er* start appearing quite frequently in spontaneous innovations around age 3, instrumental nouns with this suffix emerged a few months later, as shown in Table 24.2. Earlier, English-speaking children strongly favor root compounds made up of two nouns, as in *plánt-man* (3;0, for ‘gardener’) or *firetruck-man* (2;8, for ‘fireman’),

Table 24.2 Some innovative agent and instrument nouns

Agent nouns

- 1 D (2;5.26, as he reached across the kitchen counter): *I'm a big reacher.*
- 2 D (2;8.1, reading *In the night-kitchen*): *where's the cookers? [= cooks]*
- 3 D (3;2.17, 3;2.16, after knocking a Lego piece off a boat being built)
Mo: Oops, that one wasn't clipped on properly.
D: *I'm not a good clipper.*
- 4 D (3;2.20, poised to roll up bedside sheepskin rug): *I'm going to roll up this rug.
look out for this rug-roller! look out for this rug-roller!*
- 5 D (3;5.27): *did you know Gabriel is a stealer? sometimes he steals Judy's things.*
- 6 D (3;6.23, playing with some tiny dolls): *and these are the guarders, in case bad guys come.*
- 7 D (3;7.1, talking about being on the blue chair at preschool the week before): *guess what.
Mo. Did you make a noise?
D: I yelled and waked up the nappers. [= children who were napping]*
- 8 D (3;8.13, after story about Oliver Pig): *pigs are really good throwers.*
Mo: Oliver you mean?
D: yes, *he threw HIS ball right up into the sky.*
- 9 D (3;10.26): *someone has been messing up my room.*
Fa: Well, what's-their-names were here earlier, wh-
D (grin): *the messers?*
Fa: No, D. The carpenters.

Instrument nouns

- 10 D (3;2.28, playing with Lego): *hey! this is another kind of plane.
it has two shooters. [= guns]*
 - 11 D (3;3.3, after putting staples into a piece of paper): *take these staplers out of here! take these
staplers out of here.*
Mo: Staple?
D: *staples, take these staples out of here.*
 - 12 D (3;4.8, adding Lego blocks to a Lego wall around a field with a Lego donkey in it)
these are lockers which lock the donkey up in the farmyard.
 - 13 D (3;6.22, smoothing out a piece of paper): *I have a stamper for making things smooth.*
Mo: What's a stamper?
D: *it's something to stamp things on. [= to stamp on things]*
 - 14 D (3;9.4, looking out at the balcony where he'd been using pliers the day before):
oh, my pinchers are outside.
 - 15 D (3;9.17, talking about someone's glasses that broke): *it has only one (gesture to ear)—holder.*
 - 16 D (3;9.18, looking at a deep valance on the coffee-shop window at the airport):
Mummy, why do they have that big shader?
 - 17 D (3;9.19, watching as a mechanic placed two chocks by the plane wheels):
why did he put two loggers?
Mo: Oh, they're called chocks, and they keep the wheels from moving.
D: *why did he put the chocks?*
 - 18 D (3;9.29, taking about the paper-punch Mo had been using):
this isn't real, the puncher. it isn't a REAL toy.
-

in their spontaneous innovations and in elicitation tasks for both agent and instrument nouns. But around age 4, in elicited forms, they shift to reliance on the *-er* suffix, first for agent nouns, and somewhat later for instrument nouns (see Clark and Hecht 1982).

How soon children begin to rely on affixes for new-word formation depends in part on the typology of the language. Children acquiring French start to produce a variety of affixes in their lexical innovations between age 2 and 3 (e.g. Aimard 1975). The same holds for Hebrew, Russian, and Polish, where children make early use of a variety of affixes in their coinages (e.g. Gvozdev 1961, Chmura-Klekotowa 1964, 1971, Berman and Sagi 1981). That is, derivational affixes are favored in some languages, while compounding is favored in others. Over and above that, adult speakers prefer some options over others within a language as they coin new words. These choices can also change over time (see Schultink 1961, Dubois 1962, Chmura-Klekotowa 1970, Adams 1973).

Children also rely on the suffix *-er* in synthetic compounds in English, but they have some difficulty in learning just how to construct such compound forms. The difficulty here appears to result from the mismatch in word order for the verb and for the noun denoting the object-affected in predicates, compared to in synthetic compounds. Compare *He builds walls* with *a wall-builder*. Children learn first to attach the suffix *-er* to the appropriate core element (the verb), and produce forms like the *driver-wagon*, or the *builder-wall*, before they master the conventional word-order for the *wagon-driver* or the *wall-builder* (Clark et al. 1986). Children acquiring other Germanic languages like Dutch and German also seem to learn where to place affixes first, and only later master the word order within such compounds, and they make similar errors in word order (Clark 1993). Interestingly, the same children who make word order errors in production consistently treat the second element in such compounds, in comprehension, as the head noun. So, when given a compound like *ball-hitter*, they readily gloss it as meaning “someone who hits balls,” but if given a form modeled on one of their own productions, for example, a *throw-button* (based on a *pull-wagon*, a form elicited from the child for “someone who pulls wagons”), they still take the second noun as the head and therefore gloss it as referring to “some kind of button” (Clark 1984). With instrument nouns, 4- and 5-year-olds make similar errors in production, producing innovative forms like *fix-bike*, followed next by *fixer-bike*, for adult *bike-fixer*.

In English, *-er* is one of the earlier affixes acquired. Other affixes for constructing nominals like *-ness* and *-ity*, adjectival affixes like *-ous*, *-al*, and *-ic*, and verbal suffixes like *-ize* and *-ify*, all appear only considerably later in children’s speech, and only occasionally appear in spontaneous coinages. For the most part, children appear to start learning these Latinate suffixes and identifying the meanings they add to the roots they attach to only after they begin to read (see, e.g., Tyler and Nagy 1989, Lewis and Windsor 1996, Deacon and Bryant 2005, 2006).

How soon do children master agent and instrument forms in other languages? Children acquiring Icelandic appear to go through very similar stages to those acquiring English: after an initial stage where root compounds dominate, children begin to add the agentive suffix *-ari* to verb roots instead, and by age 4, make this their dominant

choice in elicitation tasks (Mulford 1983). Instrument nouns in *-ari* come in a bit later, in part because of competition with root compounds and conversion forms that are also frequently used for instrument nouns in Icelandic.

In French children produce innovative agent nouns with the suffix *-eur* at age 3 or even earlier. Indeed, they add *-eur* to both noun and verb roots, and among verbs often make errors early on in the form the root should take. Among the examples cited in various diary records, for instance, are *le cerceaunier* (3;0) for 'the one who mends hoops' (from the noun *cerceau* 'hoop'), *le dormeur* (3;8.14) for 'the sleeper/the one who sleeps' (from the verb *dormir* 'to sleep'), *le crêmeur* (3;8.11) for 'the eater of cream' (from the noun *crème*), *un salisseur* (3;11) for 'someone who dirties things' (from the verb *salir* 'to dirty'), *le fermeur* (5;6.10) for 'the one who shuts things' (from the verb *fermer* 'to shut'), or *un montagueur* (4;6) for 'someone who likes climbing' (from the noun *montagne* 'mountain'; cf. adult *alpiniste*). In fact, among the suffixes used in spontaneous innovations in French, agentive *-eur* accounts for some 40% of children's innovative nouns (Clark 1993). In an elicitation study (based on Clark and Hecht 1982) with slightly older children, Seidler (1988) found that children between 4;6 and 6;8 consistently relied on *-eur* in constructing both novel agent and novel instrument nouns over 90% of the time. While their choices of verb stem for these coinages identified the appropriate stem over 80% of the time with class-1 verbs, they identified the correct stem only about 55% of the time for class-2 verbs and between 30% (instruments) and 50% (agents) of the time for class-3 verbs. So one difficulty children face in forming new agent and instrument nouns in a Romance language is identifying the appropriate stem for the innovative word (see also Lo Duca 1990, for Italian).

Finally, in Hebrew, for both agent and instrument nouns, some patterns in word formation are more productive than others. Children attend to relative productivity in adult speech and appear to identify first those affixes and word-formation patterns that are more productive. They readily interpret new agent and instrument nouns, but favor one suffix overall, *-an*, when they produce both new agent and instrument forms. And, just as with other languages, their comprehension of new forms runs ahead of their own ability to produce them (Clark and Berman 1984).

24.3.2.3 Prefixes Come Later

In word formation as in inflectional development, children attend to and make use of suffixes before they begin to use prefixes (Kuczaj 1979, Hupp et al. 2009). In languages with a variety of derivational affixes, children usually master several suffixes before they make much use of any prefixes. And their first coinages with prefixes tend to be negative verbs, verbs for talking about the reversal of an action. For example, English-speaking children start to use *un-* spontaneously for reversal around age 3. A selection of such uses appears in Table 24.3. But it typically takes them several years to identify the conditions on the use of *un-* in terms of the categories of actions that can be undone or reversed, compared to those that cannot. Among the latter are actions like *break*, *run*, *scratch*, or *love*. They also have to keep track of which reversal verb forms rely on negative particles like *out* or *off*, rather than the prefix *un-*, and which have suppletive reversal forms, as in *bury* ~ *dig up*, or *hide* ~ *find* (Clark et al. 1995).

Table 24.3 Using *un-* as a verbal prefix to talk about reversal

-
- 1 D (3;0.4, as climbed over Mo's lap): *oh, your hair is unpinning.*
 - 2 D (3;1.5, after Mo proposed that they go upstairs and D started to take his blocks off the bedside table and put them into a bag for carrying up): *first I unbuild it, okay?*
 - 3 D (3;1.12, looking at pencil marks on a page that a family friend was proofreading):
are you going to ungrow these? [= erase]
 - 4 D (3;4.3, at the table, D was filling a sock with nuts from his tin in a brief spell of pretending to be Santa Claus, then he switched roles to become the recipient): *I don't know what's in my stocking.
(feels it) I'll have to unhang it.*
 - 5 D (3;5.9, to Mo): *show me how you uncatch your necklace.*
 - 6 D (3;8.1, of his pyjama top, to Fa): *can you untake this off?*
-

Table 24.4 Innovative reversal verbs in French and German

	Child coinage	Meaning	Adult form
French	dégrandir (2;6)	get smaller	rétrécir
	désendormir (3;3)	wake up	réveiller
	démonter (3;5)	come/go down	descendre
	déchauffer (3;5)	make less hot	refraîchir
	défroidir (3;6)	make less cold	rechauffer
	défatiguer (3;6)	make less tired	se restaurer, se reposer
	débâtir (3;10)	demolish	démolir
	déprisonner (4;1)	free, let go	libérer
German	forträdeln (2;1)	wheel away	fortgefahren
	abzangen (2;3)	remove with tongs	(no conventional adult form)
	aussinden (3;4)	turn off	auslöschen
	verknien (3;5)	kneel too long	sich verknien
	vergeben (5;1)	give x to the wrong person	irren im Geben
	verlaufen (5;7)	run too much	sich verlaufen
	sich verklingeln (6;3)	ring the wrong doorbell	(no conventional adult form)

Source: Based on Clark (1981).

Children acquiring English and German typically rely on a general-purpose verb like *open* or *aufmachen* in their earliest attempts to talk about reversals. These are then replaced by reliance on negative particles and, in English, also by the prefix *un-*. Children acquiring French pick up *dé-* as a reversal prefix fairly early, before age 3, and use it in such forms as *débâtir* 'unbuild' and *dégrandir* '[get] unbig = shrink'. But German

lacks a reversal prefix and relies instead on particles like *aus* and *ab* to express a reversal meaning, with occasional uses of the prefix *ver-* to convey that an action was mistaken in some way. Table 24.4 gives some typical examples of children's reversal forms in French compared to German (see Clark 1981).

24.3.3 Derivational Affixes in other Languages

Children acquire derivational affixes starting at different ages in different languages. Their first uses of affixes tend to be unanalyzed uses, where the affix is simply part of the word. So acquisition is better assessed by looking at productive uses where children coin new words with an affix, to fill some current gap in their lexicon. Such coinages may be illegitimate from the point of view of the adult's conventional lexicon because there is a word for the particular meaning intended. Such coinages are fairly soon pre-empted by the conventional word. But in assessing what children know about word structure and affix meaning, such coinages are invaluable because they demonstrate what children have analyzed so far.

For some languages, there are extensive observations on what children do and do not know at particular ages. For Slavic, there are observational and structured elicitation studies in Polish (e.g. Chmura-Klekotowa 1970, 1971, Haman et al. 2009), and in Russian (e.g. Gvozdev 1961, Ushakova 1969, 1976, El'konin 1973). For Romance, there are mainly observational studies of French (e.g. Descoeudres 1922, Grégoire 1947, Méresse-Polaert 1969, Aimard 1975, Seidler 1988); Spanish (e.g. Kernan and Blount 1966, Montes Giraldo 1976); Portuguese (e.g. Costa 1976, Mediano 1976, Figueira 1984); and Italian (e.g. Lo Duca 1990). For Germanic, there are both observational and elicitation studies of several languages, including Dutch (e.g. de Vooy 1916, Tinbergen 1919, Kaper 1959, Schaerlakens 1980, Elbers 1988); German (e.g. Stern and Stern 1928, Augst et al. 1977, Panagl 1977, Brekle 1978); Icelandic (e.g. Mulford 1983), and Swedish (e.g. Gustafsson 1979, Söderbergh 1979). Finally, there are fairly extensive observational studies and some elicitation in Finnish (e.g. Bertram et al. 2000c) and Hungarian (e.g. MacWhinney 1978), as well as in Hebrew (e.g. Berman and Sagi 1981, Berman 1982, Clark and Berman 1984).

Even for closely related languages, there are differences in how readily children can discern and analyze root vs. affix. This depends in part on phonology, which can obscure or make salient the morpheme boundaries within words. Languages also differ in how much they favor derivation over compounding, or whether they rely exclusively on one or the other. And they differ in just what is productive for contemporary (adult) speakers.

24.4 COMPREHENSION VERSUS PRODUCTION

In word-formation, one persistent finding is that children understand and offer appropriate glosses for word-forms before they are able to construct those target forms for themselves in coining new words. This is evident in the ability of 3-year-olds to provide good glosses for novel agent and instrument nouns in *-er* a good year before they start

to produce novel forms in V-*er* when asked (Clark and Hecht 1982; see also Lewis and Windsor 1996, for data from rather older children).

Indeed, when 3- to 6-year-olds were given a series of novel synthetic compounds to gloss, they were highly successful in identifying the head (final) noun as the element that identified the kind of category being talked about. When the same children were next given definitions and asked to come up with new words for the relevant agents and instruments, they produced many word-order errors, typically placing what should have been the head in initial, instead of final, position. When typical errors from each child were used as templates for a further set of novel word-forms for that child's interpretation, children consistently chose the second or final element in these compounds as the head (Clark 1984). In short, their comprehension was well ahead of their production, and the discrepancies between comprehension and production revealed both quantitative and qualitative differences.

24.5 FACTORS AFFECTING THE CONSTRUCTION OF NEW WORD-FORMS

Two factors appear to be at work in children's earliest coinages: *simplicity of form* and *transparency of meaning*. Simplicity of form captures the fact that children's earliest innovative word forms usually involve minimal changes to the root or stem form. For example, young children acquiring English or German rely on conversion before they make use of any derivational affixes. (This applies, of course, only to languages that license conversion.) It suggests that children initially prefer to make the fewest possible changes in form. At the same time, they rely heavily on known meanings—stems and affixes that are semantically transparent—in their construction of new word forms (e.g. Clark 1982, Clark and Hecht 1982).

Since languages differ in the word-formation options that are available and productive, children typically follow rather different routes from one language to another as they master the derivational options available for coining new words. Simplicity, transparency, and typology all play a role here, but a further factor, maybe the most important overall, is the (relative) productivity of each option for expressing a specific meaning-type. Adult speakers typically favor some options over others, e.g. English *-er* and Hebrew *-an* for expressions of agency, and these will therefore appear more frequently in coinages. But such preferences also change over time (e.g. Dubois 1962, Chmura-Klekotowa 1970, Guilbert 1975).

24.5.1 Productivity in Acquisition and Language Use

Productivity can be characterized in terms of formal constraints on affixes—which ones can be added where (e.g. Aronoff and Schvaneveldt 1978), and the ordering of

multiple affixes on the same root. The focus here is primarily on *word-form* rather than on the meaning being conveyed (Cutler 1980, Lieber 1980). The fewer the constraints on the form-type, the more frequently speakers are likely to use both types and tokens in the lexicon. Frequency is important for acquisition because it is correlated with children's exposure to particular word-form types (and tokens) and hence to the kinds of meanings each type expresses. But this approach does not readily distinguish between what is already "in" the lexicon, and what constitutes an "innovative" use constructed on a particular occasion to express a particular meaning.

Another view of productivity comes from normative or prescriptive accounts. These measures typically stem from the recommendations of an Academy about lexical usage and, in particular, about new-word uses in a language. And recommended word-types for lexical innovations may be disseminated through media and school systems, as in Israel, France, Sweden, among others (see Clark 1993).

Natural productivity represents a third perspective (Clark and Berman 1984). This measure is based directly on the lexical innovations adults produce (see also Plag 1999, Bauer 2001). It tracks the patterns and word-formation options currently favored by the speakers of a language. This view of productivity does not necessarily coincide with normative productivity since speakers may be unaware of, or disregard, prescriptive recommendations. Rather, it reflects actual current usage by the speakers of a language. This conforms to Schultink's (1961) definition:

Onder produktiviteit als morfologisch fenomeen verstaan we dan de voor taalgebruikers bestaande mogelijkheid... onopzettelijk een in principe niet telbaar aantal nieuwe formaties te vormen. (By productivity as a morphological phenomenon, we therefore understand language users' option... of coining an in principle uncountable number of new [word-]formations...) (Schultink 1961: 113)

While adult coinages generally observe structural constraints on word-forms, the meanings they express with these lexical innovations may fill gaps or add finer specifications to the form-meaning combinations already available in the conventional lexicon. But there is no *a priori* limit on possible lexical innovations (see also Clark and Clark 1979). Note also that speakers do not necessarily rely on the least structurally constrained form-types more often than more constrained ones. Finally, as already noted, the preferences captured by natural productivity can change over time, but they leave traces from the past in the conventional lexicon. The frequencies of derivational patterns in everyday speech may therefore diverge from the frequencies calculated from corpora based on dictionaries or word lists. The latter are best considered as a cumulative record of productivity over time, rather than a reflection of speakers' current preferences. For current (natural) productivity, one needs data on preferences by coinage-type for specific kinds of meanings.

Frequency matters here because children are attentive to it and track the frequency of forms in the language around them. This applies just as much to derivational affixes as to word stems, as shown, for instance, by children's acquisition of agentive and instrumental affixes in Hebrew (see Clark and Berman 1984). Children are also

sensitive to the relative productivity of affixes with closely related meanings, and they acquire the most productive affixes—the most frequent and typically the least constrained structurally—before less productive ones in the same domain. For example, children master English agentive *-er* before *-ist*, and they master both of these before *-ian* (Clark and Cohen 1984). More productive affixes appear to be easier to acquire: they are more frequent, more easily analyzed for meaning and form, and more readily retrieved from memory. But this can vary with just what meaning the child (or adult) is trying to convey.

24.6 SUMMARY

Overall, there are some clear cross-linguistic similarities in children's acquisition of derivational affixes for use in new words. They tend to acquire suffixes before prefixes. In doing this, they first identify the meanings of a number of suffixes and start to make use of them before they acquire any prefixes. This may well be because the status of a suffix is easier to establish in relation to word stems or roots than is the status of a prefix. Because words are processed from beginning to end, it is easier to identify words with suffixal material as words (Kuczaj 1979, Cutler et al. 1985, Cole et al. 1989, Lahiri and Marslen-Wilson 1991, Hupp et al. 2009). This could delay children's identification of prefixes as forms added to a stem, and so also delay children's identification of their meanings.

Children rely on simple options early on, with conversion when it is available. But this is used only for certain types of coinage. For example, in English, making new verbs from nouns, or in Hebrew, making new nouns from verbs (Clark 1982, 1993, Bolozky 1999). At the same time, children exhibit a strong tendency to rely on known elements as they construct new word-forms for new meanings. As a result, they go with transparency in their coinages. In doing this, they also tend to favor the most productive option available with the requisite meaning.

PART II

CHAPTER 25

INDO-EUROPEAN

PINGALI SAILAJA

THIS chapter presents some of the important derivational processes as seen in different languages of Indo-European. After a discussion of the different language groups within the larger family, it moves on to providing an overview of the more prominent word-formation processes in Indo-European. It then takes up Indo-Aryan for a more detailed analysis, with particular focus on Hindi.

25.1 THE INDO-EUROPEAN LANGUAGE FAMILY

Of all the language families known to us, the Indo-European (IE) language family has the widest presence, and the largest number of speakers across the world; it is also among the best-studied ones, globally. The most ancient known language is Hittite, about 4,000 years old; ancient Greek has a documented history of 3,500 years. Sanskrit and Latin are other ancient languages whose literature is acknowledged to be about 2,000 years old. More than 2,500 years ago, Celtic languages were spoken across most of continental Europe and into Asia Minor in the east (Figure 25.1).

According to statistics provided by *Ethnologue* (Lewis 2009), there are 426 living languages (439, including languages now extinct) in this language family, with over 2.7 billion speakers. Although the IE language family constitutes only 6.17% of the world's languages, the speakers constitute nearly 46% of the world's population. The roots of the IE language family stretch across Europe and Asia. Although the origin of the IE language family lay in the Europe-Asia axis, colonization and migration over the past few centuries have made its constituents either the first or second language of many speakers outside these regions. Consequently, languages from this family are present in almost all regions of the world, most prominently in South Asia and the Americas.

Traditionally, the language family has been divided into the Centum and the Satem sub-groups. The broad classification of the family with some examples is given in Table 25.1.



FIGURE 25.1 Indo-European groups

Source: Mallory and Adams (2006: 9).

The number of speakers for the different language groups and some specific languages are as follows. One of the older language groups, Celtic, is now spoken by around a million people in the British Isles and the Brittany region in France.

In Greece, Cyprus, Turkey, and the Balkans, around 15 million people speak Greek. Among Germanic languages, English and German have 340 and 120 million speakers respectively. Spread across Europe, the Americas and Africa, the Italic or Romance languages account for over 800 million speakers of which Spanish accounts for over 400 million and Portuguese over 200 million. In Europe, French (80 million) and Italian (70 million) are the largest. It is interesting to note that speakers in the country of origin are far outnumbered by their erstwhile colonies. Francophone Africa and South-East Asia have a significant French presence, albeit as a second language.

Slavic/Slavonic languages are present in Eastern Europe, the Balkans, Central Europe, and Northern Asia. The largest language populations are those of Russian (165 million), Polish (47 million), and Ukrainian (39 million). Baltic languages are spoken in the area proximate to the Baltic Sea in northern Europe.

Indo-Iranian is the largest branch of the Indo-European family, stretching from Europe, through the Caucasus to Bangladesh. This branch has over a billion speakers,

Table 25.1 Indo-European language family

Centum	Italic	Latin		<i>Catalan, French, Spanish, Galician, Italian, Portuguese, Provençal, Romansch, Romanian</i>
		Oscan-Umbrian		<i>Oscan-Umbrian</i>
	Germanic	West		<i>North and West Frisian, English Low Saxon—Flemish, Dutch, Afrikaans, East Frisian</i>
		East		<i>Gothic</i>
		North	Old Norse	West: <i>Icelandic, Faroese, Norwegian</i> East: <i>Danish, Swedish</i>
	Anatolian			<i>Hittite</i>
	Tocharian			<i>Tocharian</i>
	Celtic	Gaulish		<i>Gaulish</i>
		Goedelich		<i>Manx, Irish Gaelic, Scots Gaelic</i>
		Brythonic		<i>Cornish, Breton, Welsh</i>
Satem	Hellenic	Ancient Greek		<i>Modern Greek</i>
		South		<i>Bosnian, Bulgarian, Croatian, Macedonian, Serbian, Slovenian, Old Church Slavonic</i>
		West		<i>Czech, Polish, Slovak, Sorbian</i>
	Indo-Iranian	East		<i>Belarusian, Russian, Ukrainian</i>
		Dardic		<i>Kashmiri</i>
		Sanskrit		<i>Hindi, Bengali, Gujarati, Marathi, Oriya, Urdu, Rajasthani, Assamese, Nepali, Sindhi, Punjabi, Sinhalese, Romani</i>
	Iranian			<i>Old Persian, Persian, Avestan, Sogdian, Baluchi, Kurdish, Pashto, Tajik</i>
	Baltic	Baltic		<i>Old Prussian, Latvian, Lithuanian</i>
	Albanian			<i>Gheg, Tosk</i>
	Armenian			<i>Armenian</i>

across a total of 308 varieties. Within Indo-Iranian, Persian (70 million), Pashto (40 million), and Kurdish (20 million) are prominent languages. But the vast majority belongs to a sub-branch, Indo-Aryan. Led by Hindi (422 million), this group stretches east from Pakistan, and includes the majority of the populations of north India, Bangladesh, Nepal, and Sri Lanka. In the Census of India, languages such as Bhojpuri (33 million), Magahi (14 million), and Maithili (12 million) tend to be classified as Hindi. This aggregate, along with Urdu (51 million), constitutes what is called Hindustani—a bloc of over 530 million speakers. According to the Census of India, Indo-Aryan languages account

Table 25.2 Indo-Aryan—number of speakers

Language	Primary countries	Speakers in million
Asamiya	India	13
Bengali	India, Bangladesh	200
Gujarati	India	46
Kashmiri	India	5.5
Konkani	India	2.5
Marathi	India	72
Nepali	Nepal, India	15
Oriya	India	33
Punjabi	India, Pakistan	100
Sindhi	India, Pakistan	21
Sinhala	Sri Lanka	13.5

for more than 75% of the population. Table 25.2 summarizes the other major Indo-Aryan language populations.

The primary sources for the language family information are: Lewis (2009), Short (2007), and Kuriaki (2007). Clackson (2007) gives a good introduction to the topic of the structural aspects of IE. Some information is available in “Indo-Aryan languages” (2012) in *Encyclopædia Britannica*. Comprehensive data regarding the number of speakers of all Indo-Aryan languages of India are available in the Census of India (2010). The statistics related to India in all these works are based on 2001 Census information.

25.2 WORD FORMATION IN INDO-EUROPEAN

Word formation in IE languages encompasses different processes: conversion, reduplication, affixation, and compounding. However, some languages or sub-groups of languages have certain preferences. First, across all sub-groups, derivation through affixation is productive and a commonly used device for the construction of new lexical items. Compounding is another such widely used process. While reduplication is attested in most language sub-groups, it is very productive in the Indo-Aryan group. Conversion is also present in IE but to a lesser degree than derivation through affixation. English is a good example of a language that exemplifies the processes mentioned above. It must be remembered though that all of these are not equally productive.

- (1) English:
 Compounding: godson, time-frame, sunset
 Affixation: godly, sainthood, majority
 Conversion: insult_{N,V}, content_{N,A}, protest_{N,V}
 Reduplication: goody goody, itsy-bitsy, mumbo-jumbo

Vriddhi or vowel gradation is a derivational device used in Sanskrit quite productively. Hindi and other Indo-Aryan languages exhibit the same phenomenon in words of Sanskrit origin. The process is also used to some extent in Iranian, and to a lesser degree in languages like Greek.

(2)	Sanskrit:	deva-	'god'	daiva-	'divine'
		varuṇa:	'Varuna'	va:ruṇa-	'belonging to Varuna'
	Greek:	énos	'year'	é:nis	'of one year'
		oís	'ram'	ó:ia	'sheepskin'

(Adapted from Lazzeroni 1998: 121)

25.2.1 Compounding

Compounding is a highly productive process and a wide range of compounds is attested in all sub-branches of IE. All major categories are derived through compounding. Endocentric, exocentric, and dvandva or co-compounds are all possible to varying degrees in IE languages. Examples of compounds are given below. In general, NN compounds are common, followed by AN compounds. Verb compounds are less common.

- (3) Czech: N+N
 sebe-vražda 'suicide'
 self murder
 čtvrt-hodina 'quarter of an hour'
 quarter hour

(Adapted from Janda and Townsend 2000: 55)

- Icelandic:
 N+N
 stôr-hysi 'a large house'
 large house
 sjúkra-hús 'hospital'
 sickness houses
 A +N
 naut-sterkur 'ox-strong, strong as an ox'
 ox strong

raud-hoerdur 'red-haired'
red haired

(Adapted from Thráinsson 1994: 166)

Swedish: N+A
blod-röd 'blood red'
blood red

(Andersson 1994: 277)

Danish:
Adv+A
vel-smagende 'tasty' (lit. 'well tasting')
well-flavoured
A+N
små-kager 'cookies' (lit. 'little cakes')
small cakes

(Adapted from Haberland 1994: 347)

Dutch:
Adv+A
aarts-lelijk 'very ugly'
very ugly
A+A
geel-groen 'yellowish green'
yellow green

(Adapted from De Schutter 1994: 454)

Exocentric compounds are illustrated in (4):

- (4) Portuguese:
guarda-roupa 'piece of furniture to keep clothes'
keeps-clothing

Romanian:
pierde-vară 'dawdler, slowcoach'
loses-summer

(Štekauer et al. 2012: 84)

Finally, dvandva or co-compounds are illustrated in (5). Dvandvas are more common in the Indo-Aryan sub-group than in other sub-branches of IE. But they do exist in other languages as illustrated below:

Portuguese: V	nominal class marker		N	
cort-	'cut'	e	corte	'cut'
vo-	'fly'	o	voo	'fly'
reg-	'water'	a	rega	'water'

(Adapted from Rodrigues 2009: 70)

French:	V		N	
	devoir	'to have to'	le devoir	'duty'
	pouvoir	'to be able to'	le pouvoir	'power'
	savoir	'to know'	le savoir	'knowledge'

(Bauer 2003: 37)

A wide range of meanings for derived nouns is seen in the different languages. An illustration of this range of deverbal nouns from Portuguese is given in (8):

(8) Portuguese:

Event	apanha	'reaping'
	monda	'weeding'
State	zanga	'anger'
	amuo	'sulkiness'
Result	abordo	'boarding'
	branqueio	'whitening'
Concrete product	remendo	'darn'
	rasgo	'rip'
Place	abrigo	'shelter'
	assento	'seat'
Agent	pilho	'thief'
	esfervelho	'trickster'
Instrument	lixa	'sandpaper'
	aguço	'pencil sharpener'

(Adapted from Rodrigues 2009: 71)

Mostly, nouns and verbs are derived in many languages. But other categories like adjectives are also involved in the process of conversion. Nouns are derived from adjectives more frequently than the reverse process in French (Wise 1997: 108). The French examples in (9) are adapted from Wise (1997: 107–8).

(9)

	N	V
English:	pocket	pocket
	film	film
	access	access

French:	N		V	
	masque	'mask'	masquer	'conceal'
	A		N	
French:	calme	'calm'	du calme	'calm'
	responsables	'responsible'	les responsables	'person in charge'
	riches	'rich'	les riches	'the rich'
	pauvres	'poor'	les pauvres	'the poor'
German:	gut	'good'	Das Gut	'the good'
	A		V	
English:	dry		dry	
	green		green	
	V		A	
French:	aveugler	'to blind'	aveugler	'blind'
	N		A	
French:	orange	'orange'	orange	'orange'
	Adv		N	
French:	pourquoi	'why'	le pourquoi	'reason'
	comment	'how'	le comment	'the how'
	A		Adv	
French:	faux	'false'	faux	'out of tune'

While major categories may be expected and seen in the process of conversion, sometimes function words are also involved in the process of conversion.

- (10)
- | | | | | |
|----------|--------|-----------|-----------|----------------|
| | P | | V | |
| English: | up | | up | |
| | P | | N | |
| French: | pour | 'for' | le pour | 'advantage' |
| | contre | 'against' | le contre | 'disadvantage' |

25.2.4 Derivation through Affixation

Among derivational processes suffixation is the preferred option across all of Indo-European. However, some languages do attest productive prefixal derivation. For example, in Latin the following exist:

(11) Latin: Prefixation

scrībō	‘write’
a-scrībō	‘add/state in writing’
circum-scrībō	‘abridge’
de-scrībō	‘copy’, ‘transcribe’, ‘describe’
con-scrībō	‘compose’
ex-scrībō	‘copy’, ‘write out’
in-scrībō	‘inscribe’
inter-scrībō	‘write between’
prae-scrībō	‘order, direct’
pro-scrībō	‘announce’, ‘advertise’
per-scrībō	‘write in detail’
re-scrībō	‘reply’
sub-scrībō	‘write below’
trans-scrībō	‘copy, forge’

(Adapted from Štekauer et al. 2012: 143–4)

Infixation is not common, but is possible in some languages of IE. The following are examples from Spanish:

(12) Spanish: Infixation

lejos ‘far’	lej-ísim-os	‘very far’
	far-infix.aug-adv	
cerca ‘close’	cerqu-it-a	‘very close’
	close-infix.dim-.adv	

(Štekauer and Valera 2007: 196)

Circumfixation is seen in some languages. Some examples are:

(13) Dutch:

ge-t			
boef	‘rascal’	geboefte	‘riff raff’
berg	‘mountain’	gebergte	‘the mountains’

(De Schutter 1994: 454)

(Bauer cited in Lieber 1992: 155)

Suffixation is extremely common and is illustrated in greater detail in Section 25.2.5.

25.2.5 Derived Categories

All major categories are derived through affixation: nouns, verbs, adjectives, and adverbs. There are ample examples from different languages exemplifying a wide semantic range.

25.2.5.1 Nominalizations

Nominalization is a common derivational effect. Most languages have a number of suffixes that derive nominals. In general it appears that expansion of the category of noun is acceptable. Nouns are also known to be the most borrowed items in all languages. Among derived nominals, agentives are the most commonly seen in IE languages (Beard 1981a). Often the same suffix has different meanings associated with it. Agent and instrument are expressed by the same suffix in several languages.

(14)	English:	teach-er		Agent
		cook-er		Instrument
	German:	Lehr-er	'teacher'	Agent
		lehren-teach		
		Schläg-er	'whisk', 'egg beater'	Instrument
	Greek:	beat		
		elefthero-tis	'liberator'	Agent
		(eleftherono(v) 'liberate')		
	Italian:	koft-tis	'cutter'	Instrument
		(kovo(v) 'cut')		
		lavora-tore	'worker'	Agent
	Spanish:	work		
		frulla-tore	'blender'	Instrument
		whisk		
	Slovak:	trabaja-dor	'worker'	Agent
		work		
		desperta-dor '(lit.) waker'		
		wake	'alarm clock'	Instrument
		hrá-č	'player'	Agent
		play		
		počíta-č	'computer'	Instrument
		compute		

The patient meaning is attested along with the agent in the following example:

(15)	Serbo-Croatian:	propoved-nik	'preacher'	Agent
		preach		
		ranje-nik	'wounded person'	Patient
		wound		

(Adapted from Štekauer et al. 2012: 173–5)

Other types of nominalizations such as abstract and action nominalizations, and concrete nouns are also quite common in all sub-groups of IE.

	V		N		
(16) French:	animé	'animate'	anim-ation	'animation'	action/abstract
	marier	'marry'	mari-age	'marriage'	abstract nominal
	consent	'agree'	consente-ment	'consent'	abstract nominal
(Adapted from Wise 1997: 107)					
German:	verachten	'scorn'	Verachtung	'scorn'	abstract nominal
	steh	'stand'	Stadel	'barn'	concrete noun
(Ramat 1998a: 409)					

25.2.5.2 Other Categories

Other categories are also derived through affixation, primarily through suffixation. Different kinds of adjectives derived from nouns and verbs are illustrated in (17):

	N		A	
(17) English:	scholar		scholarly	
	girl		girlish	
	grace		graceful	
French:	catastrophe	'catastrophe'	catastrophique	'catastrophic'
	banque	'bank'	bancaire	'banking'
(Wise 1997: 107)				
German:	Öl	'oil'	ölig	'oily'
	Trans	'blubber'	transig	'blubbery'
(Plank 2010: 87)				
Serbo-Croatian:	jabuč	'apple'	jabučast-	'applish', 'apple-like'
(Beard 1981a: 79)				
Armenian:	mayr	'mother'	mayrakan	'maternal'
	arewelk'	'East'	arewel-ean	'eastern'
(Ajello 1998: 213)				
	V		A	
Catalan:	eixorda	'deafen'	eixordador	'deafening'
	enyor-ar	'miss'	enyoradis	'nostalgic'
(Adapted from Wheeler 1998: 205)				
French:	mang-er	'eat'	mangeable	'edible'
German:	ess-en	'eat'	essbar	'edible'
	glaub-en	'believe'	glaubwürdig	'believable'

Verbalizing suffixes exist but not to the same degree as nominalizing or even adjectival ones. In English, *-ify*, and *-ize* derive verbs as in (18):

(18)	N		V	
English:	null		nullify	
	agony		agonize	
	critic		criticize	
French:	alcool	'alcohol'	alcool-is-er	'alcoholize'
	code	'code'	codifi-er	'codify'
	A		V	
English:	solid		solidify	
French:	national	'national'	national-is-er	'nationalize'
	solide	'solid'	solid-ifi-er	'solidify'

To some extent, prefixes also derive verbs in English. Bauer et al. (2013) find some amount of productivity with the prefix *en-*.

(19)	A	V
	rich	enrich
	N	V
	power	empower
	trust	entrust

In Catalan too verbs are derived through prefixation.

(20)	Catalan:	N		V	
		pressa	'haste'	a-pressar	'hurry'
		butxaca	'pocket'	em-butxacar	'pocket'
		punt	'point'	es-puntar	'break the point of'
				remove-point	
		A		V	
		brut	'dirty'	em-brut-ar	'fouling'
		profund	'deep'	a-profund-ir	'deepen'
		malaltir	'sick'	em-malalt-ir	'fall ill'

(Adapted from Wheeler 1998: 205)

(21)	Adverbs:	A		Adv	
	English:	slow		slowly	
	French:	facile	'easy'	facile-ment	'easily'

(Wise 1997: 107)

- Nouns indicating places are derived in Armenian through suffixation.

- Wise (1997) maintains that in French, category-maintaining suffixes are less productive than category-changing ones, but are more transparent semantically. X+ *erie* is used to mean where X is made or sold.

- The same roots can take the suffix *-ier* for the meanings of producer/seller of X.

- ### 25.2.5.3 Evaluative Morphology

Evaluative morphology, also sometimes called expressive morphology, is attested in IE. Some examples of diminutives are given below:

- (26) English: dog doggy
 kitchen kitchenette
 duck duckling

German:	Haus	'house'	Hauschen	'little house'
	Würst	'sausage'	Würstchen	'little sausage'
Armenian:	naw	'ship'	nawak	'boat'
	hayr	'father'	hayrik	'little father'
(Ajello 1998: 213)				
Italian:	casa	'house'	casetta	'little house'
	stupido	'stupid'	stupidotto	'rather stupid'
French:	cigare	'cigar'	cigarette	'cigarette'
	jardin	'garden'	jardinet	'small garden'
Hindi:	qibba	'box'	qibiya	'small box'
Hindi, Bhojpuri:	beṭi	'daughter'	biṭiya	'little girl/daughter'

Examples of augmentatives are given in (27):

(27) English:	store		megastore	
	power		superpower	
Dutch:	meester	'master'	grootmeester	'grandmaster'
	handel	'business'	groothandel	'wholesaler'
Spanish:	sombrero	'hat'	sombrerazo	'large hat'
	baja	'drop'	bajón	'sharp drop'
	codo	'elbow'	codazo	'nudge, to elbow'
	macho	'macho'	machote	'tough guy'
	ceño	'frown'	ceñudo	'baleful, supercilious'

(Adapted from Lang 1990: 91, 109)

It should be noted that the degree to which the Romance and Slavic languages exemplify this process is not seen in other sub-groups. A certain degree of productivity is ascribed to such forms in Indo-Aryan by Masica (1991). He gives the following examples from Magahi all of which carry different shades of meaning and attitude for "horse":

(28) Magahi:		
ghoṛa:	ghoṛwa:	ghoṛauwa:
horse	horse-derogatory/affectionate	horse- more affectionate

(Adapted from Masica 1991: 76)

Both diminutives and augmentatives have other meanings associated with them, over and above size aspects. They can be endearments (*biṭiya* 'little girl/daughter' in Hindi) or can be used pejoratively (*machote* of Spanish). Augmentative suffixes in Portuguese have strong overtones of pejorative meaning: *mulher* 'mother', *mulherona* 'stout woman' (Parkinson 1988: 167). Evaluative morphology is not limited to any one syntactic category but is spread across many.

25.3 INDO-ARYAN

Indo-Aryan languages, particularly Hindi, are very rich in all types of word formation processes. Compounding is very productive; so are conversion, reduplication, and derivation through affixation. This section takes up the last three of these processes and shows how they interact with one another.

25.3.1 Reduplication

Reduplication, which is really a set of processes rather than a single process, is productive in Indo-Aryan (IA) to a degree not seen in other IE sub-branches. The process is said to be borrowed from Dravidian languages where it is extremely common (Masica 1991). Abbi (1992) maintains that reduplication as a process is probably diffused from Munda languages. This is one of the many areal linguistic features that enable the classification of India as a linguistic area.

Reduplication often exemplifies special meanings. It operates at all linguistic levels: phonological, lexical, morphological, syntactic, and, according to some, even semantic (Singh 1998). The process is usually captured as a phonological one. In stem reduplication, called 'morphological' by Abbi (1992), syllables which do not have morpheme status are doubled. These have an expressive function. Some are onomatopoeic; others are more than that. Some examples of nouns are given in (29):

- (29) Hindi, Punjabi, Gujarati, Marathi, Bengali: khaṭkhaṭ 'knocking'
 Marathi, Gujarati, Punjabi: jhaṇjhaṇ 'a jingling sound'
 (Masica 1991: 78-79)

Verbs and adverbs that are similarly built are shown in (30):

- (30) Verbs
 sansana:-na: 'to produce a whizzing sound', 'to have a thrilling experience'
 stem
 hinhina:-na: 'to neigh'
 stem
 Adverbs
 gaṭgaṭ 'with a quick gulp'
 phaṭa:phaṭ 'swiftly'

Syllables without morpheme status are reduplicated in (29) and (30). There is no base with an independent existence.

Whole word reduplication can have a derivative function in that a new lexical item is derived. The sense of these forms is distributive.

- (31) *bacca:* 'boy' *bacca: bacca:* 'everyone', 'each and every child'
 gali: 'street' *gali: gali:* 'everywhere', 'each and every street'

Partial reduplication can be either echo formation or can effect a morphological or grammatical change. Usually, partial reduplication is also stem-based and one of the elements has no morpheme status. For example: *puuchtach* 'inquiry' is a combination of *pu:ch* 'ask' and *ta:ch* which has no independent meaning or morpheme status. Other examples are:

- (32) *dekh rekh* 'looking after'
 see
 dekh bha:l 'looking after'
 see
 ʈhi:kʈha:k 'in order'
 right

More examples can be found in Kachru (2006: 128). Importantly, from the morphological point of view, reduplication of this kind may be considered to be compounding.

Echo formation involves substituting the first consonant with a fixed C and reduplicating the rest of the lexical item. Masica (1991) identifies an eastern apical type and a western labial type for echo formation. The fixed C is *w* in Hindi (*ʃ* in some varieties), *b* in Punjabi, *ʈ* in Bengali. Sometimes, the fixed substitution is a CV as *bi-* in Marathi and *gi* in Dakkhani (Urdu). Sinhala in Sri Lanka uses *b*. In Marathi, the influence of Dravidian is considerable, and in Dakkhani the substitution is identical to the Dravidian *gi*. Some examples from Hindi are given in (33):

- (33) *ca:y wa:y* 'tea and such things'
 dil wil 'heart and such things'
 pya:r wya:r 'love and such things'
 daulat waulat 'wealth and such things'

This type is also extended to English words:

- (34) *pen wen* 'pen and such things'
 mærej wærej 'marriage and such things'
 saikil waikil 'bicycle and such things'

Reduplicated forms interact with derivational processes. This aspect is taken up further in Section 25.3.3.5.

25.3.2 Conversion

Before we take up affixation, we demonstrate here that there are processes of conversion in Hindi that derive verbs from nouns and nouns from verbs.

(35) noun to verb

N		V	
bya:h	'marriage'	bya:h	'marry'
tya:g	'sacrifice'	tya:g	'sacrifice'

(36) verb to noun

samajh	'understand'	samajh	'understanding'
soc	'think'	soc	'thinking,' 'thought'

Verb infinitives can also be converted to nouns (where *-na:* is an infinitive marker).

(37) V

N

cal-na:	cal-na:	'walking'
ja:-na:	ja:-na:	'going'
ma:r-na:	ma:r-na:	'hitting'

All verbs have infinitive forms except *ca:hiye* 'must' (Kachru 2006). The nouns created from infinitive verbs are action nominals. Some of the derived words also have specialized meanings as in *kha:na:* which while expressing 'the act of eating' also means 'food'.

25.3.3 Affixation

Derivation through affixation is also extremely productive in many IA languages such as Hindi, Punjabi, Bhojpuri, Oriya, etc. Surprisingly, Bengali is stated not to have productive derivational affixation. Dasgupta (2003), while admitting exceptions, considers affixational derivation to be a set of related frozen forms of an erstwhile productive process.

Prefixation and suffixation are exhibited in IA languages, while infixation is absent. Of these, suffixation is more common than prefixation. There is a great deal of commonality across the IA languages in the use of prefixes and suffixes. This commonality manifests itself not just in semantic categories, but in the actual form of prefix or suffix. The primary reason for this is that in a majority of these languages, Sanskrit and Persian have had a deep influence. Forms are borrowed from these languages and we therefore find the same affix attaching to the same bases across languages. For example:

(38)	Hindi	Bengali	Oriya	Punjabi	
	a-fubh	o-fubho	a-fubho	a-fubh	'inauspicious'

neg-auspicious				
a-hinsa	o-hinja:	a-hinsa	a-hinsa	'non-violence'
neg- violence				

a- attaches to adjectives and to nouns and retains the category of the base to which it attaches.

ap 'dis' in (39) is the prefix and *ma:n* 'honour' is the base. It attaches to nouns to derive nouns.

(39)	Hindi	Bengali	Oriya	Punjabi	
	ap-ma:n	o-p-ma:n	ap-ma:nɔ	ap-ma:n	'insult'
	dis-honour				

In (40), *ku* 'bad' is the prefix and the base is *karm* 'deed'.

(40)	Hindi	Bengali	Oriya	Punjabi	
	ku -karm	ku-kɔrmɔ	ku -karmɔ	ku -karam	'misdeed'
	deed				

All the above prefixes are Sanskrit in origin and are used in many Indo-Aryan languages. Some words are pronounced slightly differently in some languages as can be seen. Oriya and Bengali have more rounded vowels, and rather than a labio-dental, these languages have a bilabial. Likewise, the following words consisting of Sanskrit suffixes are also attested in several Indo-Aryan languages.

(41)	Hindi	Oriya	Bengali	
	a:wafyak- ta:	a:bafyak- ta:	a:bɔfyɔk- ta:	'necessity'
	necessary-ness	necessary-ness	necessary-ness	
	wifej-ta:	bifej-ta:	bifejɔ-ta:	'uniqueness'
	special-ness	special-ness	special-ness	

The suffix *-ta:* is common across languages and attaches to adjectives to derive abstract nouns. This widespread influence of Sanskrit on the lexicon and processes is seen on Dravidian languages as well. Although the phenomenon is considered to be quite ancient and limited, Kachru (2006) points out that Sanskrit affixation is now seen in technical terminology where new words are being constructed.

25.3.3.1 Prefixation in Hindi

Prefixes in Hindi derive nouns and adjectives. Prefixation is a predominantly Sanskritic phenomenon in most IA languages, indeed even in Dravidian languages. Generally too, the bases to which prefixes attach are also of Sanskrit origin. Of the twenty-two nominal prefixes listed by Kachru (2006: 112–13), twenty are of Sanskrit origin. Of the ten adjectival prefixes, five have been listed as Sanskrit. There are

many Hindi prefixes listed by Kachru (2006) and Koul (2008), along with prefixes of Persian origin. Sanskrit prefixes are illustrated in (38), (39), and (40). Persian examples are given in (42):^{1 2}

(42) Persian:

be-i:ma:n	'dishonest'	be-matlab	'meaningless'
dis-honest		dis-meaning	
na:-pasand	'dislike'	na:-insa:fi:	'injustice'
not-like		not-justice	

To a very large extent, language boundaries are maintained, that is, Sanskrit prefixes attach to Sanskrit bases and Persian prefixes attach to Persian/native bases. Yet there are exceptions. Persian *be-* attaches to the Hindi *ghar* 'house' to derive *be-ghar* 'homeless'. The Sanskrit part of the vocabulary seems to remain more closed although even here we find exceptions: *an-honi* 'impossible event' where *honi* 'happening' is a word of Hindi origin and a Sanskrit prefix attaches to it. It is the existence of such forms in which affixation happens across languages that prevents us from saying, like Dasgupta (2003), that they are frozen. These extensions demonstrate that the process is still alive and productive, regardless of the degree of productivity.

In terms of the phonological changes, assimilation rules that apply in Sanskrit tend to apply even in the languages where such prefixation takes place. For instance, the prefix *ut-* exhibits changes as illustrated in (43):

(43) ut-khanan	'excavation'
up-digging	
ut-kramaṇ	'upgradation'
up-serialization	
ud-bhav	'origin'
up-world	

In these examples, there is voicing when the base begins with a voiced consonant. This prefix attaches to nouns to derive abstract nouns.

The prefix *an* exhibits a change as follows:

(44) a-hinsa	'non-violence'
neg-violence	
an-a:wafyak	'unnecessary'
neg-necessary	

¹ Standard transcription used for Hindi vowels is /ə, a/ since there is a difference in vowel quality. But these are not used here because this transcription obscures the vowel length distinction.

² Much of the Hindi data in these sections are drawn from Kachru (2006) to a great extent and from Koul (2008) and Singh and Agnihotri (1997).

Where the base begins with a vowel, *an-* is used, and in other places, *a-* is used. However, when the prefix attaches to words of native origin, there are unexpected examples such as shown in (45):³

(45)	an-parḥ	‘illiterate’		
	neg-read			
	an-mol	‘invaluable’	contra	a-mu:lya ‘invaluable’
	neg-value			not-value
	an-kaha:	‘unstated’		
	neg-said			

The categories derived in these words are: adjectives from adjectives and nouns from nouns. The prefixes here are both category changing and category maintaining. In *an-parḥ* the base is a verb. Overall, there are fewer prefixes in IA than suffixes and this tendency matches the findings of Štekauer et al. (2012) where it is observed that suffixation is the preferred option across the languages of the world.

25.3.3.2 Suffixation in Hindi

As in all IA languages, suffixes in Hindi are much more numerous and suffixation more productive in comparison with prefixation. For several examples from Oriya, see Mishra (1975), and for Punjabi, see Malik (1995). A study of the affixes across these languages demonstrates clearly their common ancestry and the common affixes across IA languages.

Sanskrit does not predominate in suffixation. Of the seventeen suffixes deriving abstract nouns listed by Kachru (2006) for Hindi, only three are from Sanskrit and two are of Persian stock. Nouns and adjectives are formed by processes of suffixation in IA. Of these, suffixes deriving nouns are the most in number, followed by adjectives and finally, verbs, of which there are hardly any. There are no verbal prefixes either. Verbs are more productively derived by attaching verb forms to nouns or adjectives. The verbs in these constructions are called verbal operators by Agnihotri (1998) and the derived forms are called conjunct verbs by Koul (2008). The most frequently used verbs for this process are *karna* ‘do’, *ho:na* ‘to be, happen’. Some examples are:

(46)	sama:pt karna	‘finish’ (V tr)
	end	Do
	sama:pt hua	‘finish’ (V intr)
	end	happen

(Adapted from Koul 2008: 102)

³ Tulsidas, a poet of the 16th century, uses *anmil* ‘not a pair’ (Ravi Ranjan pc).

In this regard, there is some paucity in Hindi affixation and other IA languages. Even in languages like English, the number of affixes deriving verbs is somewhat limited, yet that number is far greater than what is available in IA languages.

Nominalization is quite common and different types of nominalization are seen. There are both category-maintaining and category-changing suffixes. Several suffixes derive abstract nouns mostly from verbs, and to some extent, from adjectives and nouns. In most cases, the nominal or adjectival base to which an affix attaches is a full word. However, there are examples of stems to which an affix attaches. When the base is a verb, the suffixes attach to the root, that is, an item awaiting inflectional suffixes, in order to surface as a word in the language.

(47) Deverbal nominalization

V + a:i

V		N	
laṭ	'fight'	laṭa:i	'fight'
bun	'knit'	buna:i	'knitting'
caṭ	'climb'	caṭa:i	'climb, ascent'

Here, *a:i* derives an action nominal and in the case of *caṭa:i* it is also an abstract noun. It derives an abstract noun in the following cases.

(48) Deadjectival nominalization

A + a:i

A		N	
lamb-		lamba:i	'length'
long			
ūc-		ūc a:i	'height'
high			
ni:c		ni:ca:i	'depths', 'depravation'
low			

Declinable adjectives change their shape to masculine or feminine according to the gender of the noun they are modifying. So, *lamb-* must be considered to be a stem which will take either *a:* (masculine) or *i:* (feminine) suffix to be a word. Nouns are also derived from other nouns:

(49) N + iyat

N		N	
insa:n	'human being'	insa:niyat	'humanity'
ḡaxs	'person'	ḡaxsiyat	'personhood'

Furthermore, there is a suffix that has two variants with a final *a* and a final *i* which attaches unpredictably as shown in (50):

(50) V + ot-a/i

V		N	
samajh	'understand'	samjhotā (mas)	'compromise'
ma:na	'agree'	manōti (fem)	'religious vow'
cun	'select'	cunnōti (fem)	'challenge'

There is unpredictability as to which base a suffix will attach to. The suffixes attach to both transitive and intransitive verbs. In the examples given above, it is clear that the category of the final derived word is determined by the suffix.

All nouns in Hindi carry grammatical gender. The determination of the gender of derived words is also done by the suffix, but only in some cases. While -a:w, -a:wa: and -na: derive masculine nouns, -a:i, -a:waṭ, -ni: and -a:haṭ derive feminine nouns.

(51) V + a:w (mas)

V		N	
ṭhahar	'stop'	ṭhahara:w	'stoppage'
bac	'save'	baça:w	'safety', 'rehabilitation'
lag	'attach'	laga:w	'attachment'
jam	'accumulate'	jama:w	'accumulation'
bah	'flow'	baha:w	'flow'

(52) V + a:wa: (mas)

V		N	
bul	'call'	bula:wa	'call'
chal	'cheat'	chala:wa	'illusion'
pachta	'regret'	pachta:wa	'regret'

(53) V + a:waṭ (fem) a:waṭ

V		N	
likh	'write'	likha:waṭ	'handwriting'
ban	'become'	bana:waṭ	'formation'
mil	'meet', 'mix'	mila:waṭ	'adulteration'
saja:	'decorate'	saja:waṭ	'decoration'

(54) V + ni: (fem)

V		N	
kar	'do'	karni:	'deed'
bhar	'pay'	bharni:	'payment'
kah	'say'	kahni:	'utterance'

While the gender of the noun is determined by the suffix above, some other suffixes do not carry specification with regard to gender. Each derived word must have its own gender assigned to it. This means that suffixes may either carry gender or may remain unspecified for this feature. The suffixes -an and a:n are not specified for gender and derived nouns are unpredictably masculine or feminine.

- [illegible]

The first meaning in (61) would fall into the set of patient derivatives. The second meaning has a locative sense.

Some of the words above have feminine forms as well: *lekhika:*, *fikfika:*, *pari:kfika:* (the last being very rare). However, these feminine forms are going out of use. The words *sudha:rak* and *sa:dhak* are grammatically masculine. That is, the default usage is as a masculine noun. However, if a woman is to be referred to, the sentence structure reveals it with the use of a modifier. In that sense, this derivation may actually be considered to be gender-neutral.

A very common and productive suffix in current times is *wa:la:* 'person.' Bhojpuri and Punjabi also use this suffix. In fact, Verma (2003) considers *wa:la:* in Bhojpuri to be more productive than in Hindi. Almost any verb can take the suffix *wa:la:* in Hindi to form an agentive. The feminine form is *wa:li:*. In all the verbs below the structure is that of [V + na: (infinitive) + e (inflection)] to which *wa:la:* attaches. The form to which *wa:la:* attaches is not the verb root but the inflected infinitive form (Kachru 2006: 136), called the infinitival participle by Singh and Agnihotri (1997: 61).

(62) -*wa:la:* from verbs

kha:ne-wa:la:	'one who is eating/eats'
eat-ing-person	
dekhne-wa:la:	'one who is seeing/sees'
see-ing-person	
udne-wa:la:	'one who is flying/flies'
fly-ing-person	
ja:ne-wa:la:	'one who is going/goes'
go-ing-person	

This suffix also attaches to causative forms:

(63) kha:ne-wa:la:	'one who is eating/one who eats'
khila:ne-wa:la:	'one who is feeding/feeds'
khilwa:ne-wa:la:	'one who is making someone feed/makes someone feed'

The last item is a second causative and the *wa:la:* form would be rather rare. The meanings are very transparent here. The suffix *wa:la:* also attaches to nouns where it is less productive:

(64) -*wa:la:* from nouns

du:dh-wa:la:	'milk man'
milk	
bandar-wa:la:	'monkey man'
monkey	
pagṛi-wa:la:	'turbaned man', 'person with prestige'

turban	
sa:ʃi-wa:la:	‘sari man’, ‘one who sells saris’
sari	
laɖki-wa:la:	‘person from bride’s side’
girl/bride	
beʃi-wa:la:	‘one with a daughter’
daughter	
jhopɖi-wa:la:	‘one in a hut’
hut	
rikʃa:-wa:la:	‘rickshaw puller’
rickshaw	

The suffix *-wa:la:* can attach to plural nouns also:

- (65) arma:n-ō-wa:la: ‘one with a number of desires/ambition’
hope-pl
kapʃ-e-wa:la: ‘one who deals in clothes’
cloth-pl

Not all of the examples in (62–65) are agentives. Some indicate ownership. As is evident, some also refer to professions.

The participle form participating in conversion has been noted for many languages by Štekauer et al. (2012). Here we have a derivational suffix attaching to inflected forms. *wa:la:* words can be also be adjectives: *lamba: wa:la aadmi* ‘tall man’ where *lamba:* is ‘tall’ and *aadmi* is ‘man’. The *-na:* (*calna:*) type of construction illustrated in (37) is similar to *wa:la:*. Both types are extremely productive. In both cases, that is, the *wa:la:* and the *-na:* constructions, the infinitive form of the verb is relevant for affixation or conversion.

25.3.3.4 Derivational Affixation in Relation to Other Processes

The interplay of affixes and compounding is along expected lines in Hindi. Surabhi Bharati (1994) demonstrates the relationship between the morphological processes in Hindi in a level-ordered theory. However, what is interesting is that inflected forms are seen inside compounds, reduplicated structures (which may be considered to be compounds), and some inflected forms as illustrated above (as in (65)) can also take derivational affixes. Reduplicated stems can be nominalized as below:

- (66) V + a:haʃ (fem)
- | V | | N | |
|------------|---------|----------------|------------|
| bhinbhina | ‘buzz’ | bhinbhina:haʃ | ‘buzzing’ |
| jhanjhan a | ‘clang’ | jhanjhan a:haʃ | ‘clanging’ |
| girgir a | ‘beg’ | girgira:haʃ | ‘begging’ |
| gunguna | ‘hum’ | gunguna:haʃ | ‘humming’ |

Echo forms and *wa:la:* forms are different in comparison with other derivational suffixes. In general, when echo words are formed out of derived words lexical integrity is maintained. That is, it is not possible to form an echo word with just the root without the derivational affix.

- (67) *likh-na: wikhna:* **likh wikhna* ‘writing and other things’
 write-
 infinitive
apma:n wapma:n **apwapma:n* ‘insult and such things’
 insult

(Adapted from Mohan 2006: 123)

On the other hand, when echo forms are constructed out of *wa:la:* words two variants are possible. Words of different categories are exemplified in (68).

- (68) *ni:ce-wa:la-wice:-wala:* and *ni:ce-wice:-wa:la:* ‘the one below’
 below-person
ghar-wa:la-war-wa:la: *ghar-war-wa:la:* ‘owner of a house’
 house-person
du:dh-wa:la-wu:dh-wa:la: *du:dh-wu:dh-wa:la:* ‘milkman’
 milk-person
ga:ne-wa:la-wa:ne-wa:la: *ga:ne-wa:ne-wa:la:* ‘the person singing’
 sing-inf-infl-person

(Adapted from Mohan 2006: 123)

The words in (68) demonstrate that echo-formation and derivation of *wa:la:* do not follow a strict pattern. The behavior of *wa:la:* and the base to which it attaches demonstrates clearly that the strict linearity of derivations followed by inflections is untenable for this affix at least.

Some examples of inflections inside compounds can be found in other IE languages. In English, they are generally either pluralia tantum or suppletive plurals. For example, in *arms race*, the word *arms* has a specific meaning, or *teeth marks* contains a suppletive form. Furthermore, the productivity of these structures is limited, whereas in Hindi, inflected forms feeding derivational *wa:la:* are totally productive.

This suffix *wa:la:* has a role to play in syntax as well. One of the features of this affix is that it can attach to possessive pronouns like *mera* ‘my’ as in *mera: wa:la: pink* ‘My type of pink.’ Also, the argument structure of the verb to which it attaches can remain intact and be realized in a sentence:

- (69) a *miṭha:i* *kha:ne* *wa:la:* *a:gaya*
 sweets eat-inf-infl one come-past-mas.
 ‘The one who eats sweets has come.’

- b ga:na: ga:ne wa:la: mera: beṭa: hai
 song sing-inf-infl one my son is
 'The one singing is my son.'
- c ghode becne wa:la:
 horses sell-inf-infl one
 'One selling horses'
- d lakḍi kaṭne wa:la:
 wood cut-inf-infl one
 'One chopping wood'

However, the inflection must be considered to be attached to the verb and not to the whole unit of noun plus verb since the following is possible:

- (70) hindi ga:na: ga:ne w a:la
 Hindi song sing-inf-infl one
 'One singing a Hindi song'

Masica (1991) also observes that *wa:la:* has syntactic properties. Mohanan (1995) states that *wa:la:* attaches to forms that may be classified as those involving noun incorporation. And she considers them to be constructed in the lexicon.

In terms of overall productivity of all the processes discussed here, it may be concluded that those items that are derived out of Sanskrit affixes are the least productive, prefixation in general being less productive. Of the rest, Hindi suffixes are obviously the most productive. Where an affix attaches to bases of a particular language origin, the affix is bound to be less productive.

25.3.3.5 *Influence of English*

Although there are many English words in all IA languages and the lexicons are replete with borrowings from English and other European languages, especially Portuguese, IA has not borrowed derivational forms from these languages. However, there are examples of English words taking Hindi suffixes:

- (71) phone wa:la: 'phone operator'
 paper wa:la: 'paper boy,' 'news agent'

Some examples of English words taking Hindi derivational affixes are cited by Borowiak (2007), reproduced in (72–77), but these need to be taken cautiously. What is important here is to what extent the words are assimilated into Hindi. Shapiro (2003: 275) gives the example of *kampyūṭari:karaṇ* 'computerisation.'

- (72) i: 'like'—forming adjectives from noun stems; nouns indicating profession
 film filmi: 'filmy'
 film
 d̪a:k̪ar d̪a:k̪ari: 'medicine,' 'medical science'
 doctor
 kap̪ta:n kap̪ta:ni: 'captaincy'
 captain
 afsar afsari: 'duties of an officer'
 officer
- (73) -iyat forms: abstract noun
 hain̪ṣam hain̪ṣamiyat 'handsomeness'
 afsar afsariyat 'officialdom'
- (74) -ka:r (agent nouns from nominal stems)
 film filmka:r 'film maker'
- (75) a:nkan (action nominal)
 film filma:nkan 'filming'
- (76) da:ta: (Sanskrit suffix—noun 'giver')
 woṭ woṭda:ta: 'voter'
 vote

An English word can also be prefixed with a Persian prefix:

- (77) be- (without)
 ṭikaṭ beṭikaṭ 'without a ticket'

(Adapted from Borowiak 2007: 12)

The English words are completely assimilated into Hindi. So, this type of affixation is unsurprising. But the relative newness of the forms demonstrates that derivation is a process that is quite alive.

Although we stated earlier that Hindi has not borrowed English affixes, it does not mean that English suffixes do not attach at all. A notable case is the use of *-ify*. Almost all IA languages, and even Dravidian languages, attach *-ify* to indigenous verbs and use the words in English sentences. These productively constructed words, however, are restricted to informal and colloquial use. Here, an English affix is borrowed into Hindi, indeed into most Indian languages.

- (78) ma:r-o-fai 'hit'
 hit-imp-fy
 bhaṭka:-o-fai 'instigate'
 instigate-imp-fy
 d̪ara:-o-fai 'scare'
 scare-imp-fy

Again, these are attached to inflected forms. The sole purpose of these forms is to enable them to be used in English sentences.

In conclusion we may say that Hindi exhibits a very robust lexicon with a number of derivational processes. Derivation in Hindi interacts with other word formation processes in very interesting ways which bring into question the strict linearity of word-formation processes.

CHAPTER 26

URALIC

FERENC KIEFER AND JOHANNA LAAKSO

26.1 INTRODUCTION

26.1.1 The Uralic Languages

It is practically impossible to describe derivation in the Uralic language family within one chapter. The time depth and internal diversity within the Uralic/Finno-Ugric language family are comparable to those of Indo-European. Moreover, the Uralic languages are very unevenly researched and documented; even for the smallest of them (except those which became extinct before the time of modern field research in the late 19th century) there are at least some kinds of grammars and dictionaries or word lists, but studies of specific questions such as derivation are not always available. For this reason, we will only present concise descriptions of the derivation in some selected Uralic languages, trying to ensure that all the main branches of the language family are represented.

The Uralic languages are spoken in a large area stretching from Eastern and Northern Europe (Russia, Estonia, Latvia, and Scandinavia) to Western Siberia; Hungarian, due to the half-nomadic prehistory of the Hungarians, forms a geographic outlier. The internal taxonomy of the Uralic language family is subject to some controversy (especially as to whether the nodes should represent binary or multiple splits, in particular as concerns the basic division of Uralic into Samoyedic and Finno-Ugric = non-Samoyedic), but there is a consensus that the family consists of six main branches:

1. **Finnic** (Finnish, Estonian, and some minority languages: Livonian in Latvia, Votic, Ingrian, Veps, and Karelian in Northwest Russia; the Finnish varieties Meänkieli and Kven in the Far North of Sweden and Norway, respectively, have recently been recognized as independent languages, while two southern Estonian varieties, Võro and Seto, are being developed and seeking recognition) and Sámi (a chain of sister varieties stretching from Northern Scandinavia to

the Kola Peninsula). Of these, Finnish (with *c.* 5 million speakers) and Estonian (with *c.* 1 million speakers) are full-scale nation-state languages, all other Finnic languages, their numbers of speakers ranging from more than 50,000 (Karelian) to just a few (Livonian and Votic on the verge of extinction), are more or less endangered. The Sámi dialect continuum is usually divided into ten Sámi languages, among which the greatest is Northern Sámi (spoken in Norway, Finland, and Sweden) with approximately 30,000 speakers; other Sámi varieties have many fewer speakers, some are extinct or extremely endangered.

2. Mordvin (two varieties, Erzya and Moksha, in European Russia, with more than 400,000 speakers in sum¹);
3. Mari (two standard varieties in European Russia, with almost 400,000 speakers in sum);
4. Permic: Komi (two standard varieties, with more than 160,000 speakers) and Udmurt (more than 300,000 speakers) in European Russia;
5. Ugric (Hungarian in Hungary and adjacent areas, with *c.* 13 million speakers; the Ob-Ugric languages Mansi and Khanty in Western Siberia, Khanty with almost 10,000 speakers, while Mansi is spoken by probably less than 1,000);
6. Samoyedic (Nenets, Enets, Nganasan, Selkup and some corpus (= extinct) languages in Siberia, Nenets also in the Far North of European Russia; the numbers of speakers for the Samoyedic languages still alive range between more than 20,000 for Nenets and a few dozen for Enets).

26.1.2 On the Research Tradition

The existing studies on Uralic word-formation can be roughly divided into the following types:

- a. The two standard works dealing with (the history of) derivation in Uralic as a whole: Györke (1934) and Lehtisalo (1936), followed by some general reflections by Kangasmaa-Minn (1987, 1992). Uralic comparisons concentrating on a certain derivational suffix or function have only been undertaken in a few cases (e.g. Lazar (1975) on the derivation of abstract nouns; Laakso (1997) on verbalizer suffixes).
- b. Chapters on word formation (usually fairly scanty, sometimes not very far beyond listing examples of the “famous opulence of derivatives” in the major

¹ Due to the endangered state and narrow scope of use of the minority languages of Russia, the speaker statistics, based on the most recent Russian census from 2010, are not completely reliable; hence, only very coarse estimates are given here, and the numbers should be treated with caution.

- Uralic languages) in general introductions into Uralic comparative grammar: e.g. Collinder (1960: 220–32), Majtinskaja (1974), Hajdú (1987: 257–60).
- c. Language-specific studies on word (often: verb or noun) formation in a “minor” Uralic language. Most of these have been written from a historical-comparative point of view, and/or on the basis of classical dialect material collected before World War I; see, e.g., Sauer (1967) for Khanty, Kneisl (1978) for Komi, Riese (2001) for Mansi. However, there are also some descriptive studies based on more recent fieldwork, such as Wagner-Nagy (2001) for Nganasan, or even written by native speakers, such as Ruong (1943) on Pite Sámi, Zajceva (1978) on Veps, or Markianova (1985) on Karelian.
 - d. Language-specific studies on word formation, describing a modern standard language and often with a strong emphasis on language correctness, or chapters of this type as part of a standard “academic grammar.” Some of them are theoretically oriented such as Kiefer (2003) and Kiefer (forthcoming) on Hungarian word formation.
 - e. Studies on a suffix, a suffix group, or a type of suffixes in a certain language or a sub-group of the Uralic languages.

For a more recent overview of the history of word-formation in Uralic cf Laakso (forthcoming).

Most Uralic languages of Russia (i.e. all except the Finnic group) have standard orthographies based on the Cyrillic alphabet. Here, these languages will be written as in most Uralistic sources, using the modified Uralic/Finno-Ugric transcription (a.k.a. UPA or FUPA; Posti and Itkonen 1973). The most important features of this transcription system: palatalization is marked with an acute accent (*ś*, *ń*), mid vowels with a breve under the vowel character (*e* for “back/mid e,” *i* for “back/mid i”).

26.2 GENERAL CHARACTERISTICS

In all Uralic languages, both inflectional and derivational morphology is solely or predominantly suffixal. Prefixes, if any, are clearly innovations and only used in a few exceptional cases, mainly in derivational morphology. The most productive and also the most debated case consist of the preverbs in Hungarian (Kiefer 2010). Negative prefixes for various types of adnominal negation or for forming antonyms also appear in some Uralic languages (cf. Finnish *epä*-, Estonian *eba*-, and Northern Sámi *eahpe*-, etymologically a participle form of the negative auxiliary). Many Uralic languages of Russia use the Russian negative prefix *ni*- in negated pronouns or pronominal adverbs (e.g. Mari *ni-gušto* ‘nowhere,’ Karelian *ni-ken* ‘nobody’).

The Uralic languages typically have rich and complex derivational systems, including many suffixes which can be traced back to Proto-Uralic. Characteristically, highly productive derivational suffixes (sometimes functionally close to inflectional ones)

exist side by side with less productive and obsolete ones. Furthermore, in the course of linguistic emancipation and language reforms, starting at the turn of the 19th century or later, which resulted in the modern Hungarian, Finnish, and Estonian standard languages, numerous suffixes (noun suffixes in particular) were revitalized or coined which were used in a number of neologisms (for instance: Hungarian *úszó-da* ‘swimming pool,’ *iro-da* ‘office (‘writing place’)’; Finnish *el-iö* ‘organism, living thing,’ *kolm-io* ‘triangle,’ *luk-io* ‘grammar school (‘reading place’)’) but are not necessarily productive in today’s language.

In the following survey we will attempt to restrict ourselves to productive derivational suffixes. However, it must be noted that not all available descriptions have tackled issues of productivity in a systematic way. Moreover, the speakers of all Uralic minority languages today are bilingual, and at least in the language use of the younger generations the local majority language (Russian, Norwegian, etc.) often dominates to such an extent that present-day productivity patterns are difficult to assess. A further problem is posed by the characteristic richness of *Aktionsart* suffixes for momentaneity/semelfactivity (“X once,” “X quickly”), continuity (“X for some time,” “go on X-ing”), frequentativity (“X often,” “X many times”), etc. These derivatives are prone to lexicalizations in various directions and are interconnected with aspectual meanings (for instance: frequentative meanings of “doing often” may be related with imperfective aspect, that is, “not doing properly,” “not finishing”). In the Uralic languages of Russia, these meanings may develop towards aspect proper (see e.g. Puura 2010), and the standard grammars often describe them from the viewpoint of the Russian aspect system and the Russian grammar tradition.

26.3 OVERVIEW OF DERIVATIONAL PROCESSES IN SELECTED URALIC LANGUAGES

26.3.1 Finnish

There is a plethora of studies on various aspects of derivation in Finnish, but most of them, regrettably, are only available in the Finnish language, while the standard handbooks available in other languages typically treat derivation very concisely or not at all. Comprehensive descriptions can be found in Lauri Hakulinen’s handbook (Hakulinen 1961, 1979) or in the practical handbook of Lepämaa et al. (1996). The most up-to-date comprehensive treatment of derivation in Finnish is included in the new descriptive grammar of Finnish by Auli Hakulinen et al. (2004).

26.3.1.1 Noun Derivation

Within the macro-category *nomen* (nouns, adjectives, pronouns, numerals—i.e. words that can be inflected in case and number), the boundary between adjectives and nouns

is not absolutely clear-cut, and many suffixes can be used for deriving both nouns and adjectives. Thus, for instance, derivatives in the very productive *-lAinen*,² expressing origin (e.g. *lontoo-lainen* ‘Londoner, from London’), function both as adjectives and as nouns.

Frequent and (fairly) productive suffixes for deriving nouns from nouns or adjectives are for example (i) *-(U)Us* for abstract properties (*laps-uus* ‘childhood,’ *kaune-us* ‘beauty,’ *köyh-yys* ‘poverty’); (ii) *-stO* for collective nouns (*saari-sto* ‘archipelago, group of islands’), (iii) *-lA* for localities (*sika-la* ‘pig-sty,’ *kahvi-la* ‘coffee shop,’ *mummo-la* ‘Grandma’s house’). Unlike most other Finnic and Sámi languages, Finnish has no productive diminutive derivation (*pace* Laalo 2001; his examples of “diminutives” are a semantically and morphologically heterogeneous group of irregular affective, slang, and motherese formations). Feminine equivalents for occupational terms and the like can be formed with *-tAr* (*johtaja-tar* ‘[female] director’), but this practice, introduced in the 19th century, is receding, and *-tAr*-less forms are normally used for both genders (Engelberg 2002). (Note that the Uralic languages have no grammatical genders.)

Action nouns are derived from verbs with the highly productive *-minen* (*syö-minen* ‘(the act of) eating’), beside numerous less productive action noun suffixes, the distribution of which is partly lexically, partly phonotactically determined (*kaat-o* ‘(the act of) felling,’ *juoks-u* ‘(the act of) running, run,’ *ava-us* ‘opening,’ *hakk-uu* ‘hitting, chopping,’ *juhli-nta* ‘celebration’). Some of the latter types are often lexicalized or develop towards nouns denoting results of actions (*maala-us* ‘painting,’ *pure-ma* ‘bite,’ *polk-u* ‘the act of treading’ > ‘path’). For agent nouns, the suffix *-jA* is highly productive (*opetta-ja* ‘teacher,’ *julkaisi-ja* ‘publisher,’ *hiihtä-jä* ‘skier’). Fairly productive are also *-in* for instruments (*vahvist-in* ‘amplifier,’ *tuulet-in* ‘ventilator’) and *-mO* for places of action (*teurasta-mo* ‘slaughterhouse,’ *valvo-mo* ‘monitoring room’).

26.3.1.2 Adjective Derivation

Derived adjectives can be divided into three semantic types: possessive (relational), caritive (privative), and moderative. Moderative adjectives are formed from adjectives with certain suffixes, the most productive of which is *-hkO*: *piene-hkö* ‘smallish, fairly small,’ *lihava-hko* ‘fairly fat.’ The other types are most often formed from nouns, although adjective bases can also occur.

Of the suffixes for possessive or relational adjectives, the most versatile is *-(i)nen* (which also appears as a semi-productive suffix for diminutive nouns and as part of many complex suffixes): e.g. *lika-inen* ‘dirty’ (from *lika* ‘dirt’), *kulta-inen* ‘golden’ (from *kulta* ‘gold’), *pitkä-tukka-inen* ‘long-haired’ (*pitkä tukka* ‘long hair’). Further productive suffixes are, for example, *-lAinen* (cf. above), *-mAinen* (‘resembling X’: *tyttö-mäinen* ‘girlish’), *-llinen* (‘having X’: *huole-llinen* ‘careful’). There is also a host of less productive

² Due to the vowel harmony in Finnish, Hungarian, and some other (although by far not all) Uralic languages, suffixes may have allomorphic variants with different vowels. For these cases, archiphonemic notation with capitals is used (e.g. Finnish *-lAinen* denotes the suffix variants *-lainen* and *-läinen*).

and non-productive adjective suffixes (e.g. *liha-va* ‘fat, obese’ from *liha* ‘flesh, meat,’ *luku-isa* ‘numerous’ from *luku* ‘number,’ *parra-kas* ‘bearded’ from *parta* ‘beard’ etc.).

A specific type of adjectives, denoting an equal extent of a property, is formed with *-Uinen* from adjectives, or with *-inen* from nouns denoting dimension of measurement. These adjectives require a possessor attribute (in the genitive case) expressing the standard of comparison: *kahde-n metri-n syv-yinen* ‘two metres deep’ (cf. *syvä* ‘deep’), *miehe-n pit-uinen* ‘as tall as a man,’ *norsu-n paino-inen* ‘heavy as an elephant, of an elephant’s weight’ (cf. *paino* ‘weight’).

Caritive (privative) adjectives are formed from nouns with *-tOn*: *huole-ton* ‘careless,’ *työ-tön* ‘without work, unemployed.’ Together with the nominalizer/participle suffix *-mA*, this suffix can also productively be attached to verb stems, to form negative antonyms for present participles, active or passive: *ole-ma-ton* ‘non-existent’ (antonym of *ole-va* ‘existent, being’), *korvaa-ma-ton* ‘irreplaceable’ (antonym of *korvat-tava* ‘what is/ can be replaced’).

The participles, forming a 2×2 system (active/passive, non-past/past), are functionally very close to adjectives proper. At least certain lexicalized and frequent participle forms can be interpreted as conversion from participles to adjectives: *hurmaa-va* ‘charming’ (active non-past), *naure-ttava* ‘ridiculous’ (‘to be laughed at,’ passive non-past), *väsy-nyt* ‘tired’ (active past), *kiro-ttu* ‘damned’ (passive past).

26.3.1.3 Verbal Derivation

Hakulinen et al. (2004) divide the multitude of Finnish verbal derivational suffixes into three main types on the basis of the meaning of the derived verbs; this may be the most practical solution, as many suffixes can be attached to both noun and verb stems, so that a division into denominal and deverbal derivations would not bring much clarity. These types are (i) (transitive) verbs which denote causing a change of state (“causatives” in the wide sense); (ii) (intransitive) verbs which denote (experiencing) a change of state (reflexive and transitive); and (iii) verbs in which the meaning of the base is “modified.”

However, not all verbal derivational suffixes fall neatly into just one of the categories. In particular, the so-called verbalizer suffixes (cf. Laakso 1997) whose function is simply to form verbs out of nouns (the semantics of the verb are determined by the semantics of the base and by general, encyclopedic factors) may form verbs belonging to many different semantic types. For instance, *-tA-* can form causative (*kaunis-ta-* ‘to make (something/somebody) beautiful’ from *kaunis* ‘beautiful’) or essive verbs (*sairas-ta-* ‘to be ill’ from *sairas* ‘sick, ill’).

Causative verbs are typically formed with the suffixes *-tA-* and *-ttA-* (or complex suffixes containing these suffixes): *syö-ttä-* ‘to feed’ (from *syö-* ‘to eat’), *vero-tta-* ‘to collect tax’ (from *vero* ‘tax’), *ase-ista-* ‘to arm’ (from *ase* ‘weapon’), *make-utta-* ‘to sweeten’ (from *makea* ‘sweet’). The suffix *-ttA-*, sometimes in combination with the reflexive *U*, is used for deverbal curative-causative (factitive) derivation (“make somebody X,” “have something X-ed”): *leikka-utta-* ‘to have cut’ (from *leikkaa-* ‘to cut’), *laula-tta-* ‘to make (somebody) sing’ (from *laula-* ‘to sing’). Curative-causatives are also used in a specific subjectless construction type (the “emotion causative,” see e.g. Siirainen 2005), in which

the experiencer is the object: for example *minua laula-tta-a* ‘I feel like singing’ (‘Ø makes me sing’).

Denominal causative (transitive) verbs are also formed with *-i-* and *-Oi-*, for example *soker-oi-* ‘to strew with sugar, to sugar,’ *satul-oi-* ‘to saddle.’ The suffix *-Oi-* is often used for the adaptation of loan verbs, for example *install-oi-* ‘to install,’ *edit-oi-* ‘to edit.’ Monosyllabic loan verbs in particular can also be adapted with *-AA-* (historically going back to **-tA-*): *buukk-aa-* ‘to book,’ *tsekk-aa-* ‘to check.’

Reflexive and transitive verbs (intransitive verbs denoting change of state) are formed from both nouns and verbs with the reflexive suffix *-U-* or complex suffixes containing the element *-U-* (e.g. *must-u-* ‘to become black’ from *musta* ‘black,’ *kipe-yty-* ‘to become sore’ from *kipeä* ‘sore,’ *pese-yty-* ‘to wash oneself’ from *pese-* ‘to wash,’ *sirpal-oitu-* ‘to become fragmented’ from *sirpale* ‘fragment’). From adjectives, transitive verbs are also formed with the suffix *-ne-* (e.g. *piene-ne-* ‘to become small(er)’ from *pieni* ‘small’) or with the complex suffix *-ntU-* (< *-ne-* + *-tA-* + *-U-*), for example *piene-nty-*.

A host of suffixes form deverbal verbs with frequentative (*lue-skele-* ‘to read leisurely,’ *hypp-i-* ‘to jump repeatedly’) or momentaneous (*luk-aise-* ‘to read quickly through,’ *hyp-ähtä-* ‘to make one small jump’) meanings, sometimes even combinations of them (*hyp-äht-ele-* ‘to make small jumps repeatedly’).

26.3.1.4 Adverbial Derivation

The number of suffixes for adverbial derivation is much smaller than in noun and verb derivation, and most of them are not fully productive. As in many other Uralic languages, the boundary between adverbial derivational and adverbial case suffixes is not completely clear-cut. Some adverbial suffixes resemble case suffixes: the suffix *-tse* for adverbials denoting path of transportation (*meri-tse* ‘by sea,’ *puhelimi-tse* ‘by telephone,’ *sähköposti-tse* ‘by e-mail’) is also known as the “prolative” and sometimes described as an extension to the case system. On the other hand, many suffix chains containing case and other suffixes are grammaticalizing into what must be synchronically analyzed as adverb suffixes. For example, *nuku-ks-i-ssa* ‘asleep’ consists of the verb stem *nukku-* ‘to sleep,’ a noun derivation suffix *-s* (*-ks-*), the plural suffix *-i-* and the inessive case suffix *-ssA*, but as the noun **nukus* never appears in other than the local case forms, in an adverbial function, the form cannot be interpreted as a normal inflected noun.

More productive adverbial suffixes are, for example, *-sti* (forming adverbs of manner from almost any adjective: *kaunii-sti* ‘beautifully,’ *huolellise-sti* ‘carefully’) or *-ttAin* (forming distributive adverbs: *päivi-ttäin* ‘daily,’ *pari-ttain* ‘in couples, two by two’).

26.3.1.5 Conversion

As stated above, within the macro-category *nomen* there are no absolutely clear-cut boundaries between nouns and adjectives: if their semantics allow, nouns (derived or underived) can occasionally be used as adjectives or vice versa, and some word types, for example derivatives in *-lAinen* denoting the origin of a person, can be used both as nouns and as adjectives. This N/A ambiguity is usually not considered conversion or “zero derivation” in the Finnish grammar tradition.

Between nouns and verbs there is no true productive conversion, although there are some underived N/V word pairs with identical stems (*sylke*- ‘to spit/spittle’) and some further word pairs with identical suffixes (*paini* ‘wrestling’ / *paini*- ‘to wrestle’, both historically derivatives in *-i*-).

26.3.1.6 Other Types of Word Formation

Back-formation and blending are not very productive but do occur, as in *tarra* ‘sticker; velcro’ (from *tarrata* ‘to clutch, to get stuck’) or *silla* ‘rayon, fabric made of cellulose fiber’ (from *selluloosa* ‘cellulose’ + *villa* ‘wool’). Clipping, complex clipping in particular, is used especially to form the first, attributive part of a (new) compound term: *luomu*- (from *luonnonmukainen* ‘natural, ecologically produced,’ e.g. in *luomu-viljely* ‘eco-agriculture,’ *luomu-maito* ‘eco-milk’), *luki*- (from *lukemis- ja kirjoitus*- ‘reading and writing,’ e.g. in *luki-häiriö* ‘reading and writing disorder, dyslexia’).

Affective and slang words, nouns in particular, are often formed irregularly and analogically, by way of holistic “word *Gestalt* models” for which Hakulinen et al. (2004) use the term *muotti* ‘mould’: the base word stem may be truncated or phonologically modified, after which a “quasi-suffix” can be attached to it: *kylpyhuone* ‘bathroom’ > *kylpp-äri*, *nurmikko* ‘lawn’ > *nur-tsi*, *bensiini* ‘petrol, gasoline’ > *bens-a*, *häiriö* ‘disorder’ > *hii-kkä*.

26.3.1.7 Some Remarks on Estonian and Sámi

Estonian is closely related to Finnish, although the languages are not quite mutually intelligible, and most of the derivational suffixes of Finnish have cognates in Estonian. However, their functions, distributions, and productivity display major differences. For instance, the verb suffix (*t*)*tA* does appear in Estonian but it does not form (productive) deverbal curative-causative (factitive) verbs.

Estonian, phonologically less archaic than Finnish, has undergone a typological shift towards increasing fusion and loss of syntheticity (although there are also agglutinative cross-currents; cf. Grünthal 2000). Accordingly, the derivational system in Estonian is more compact, that is, there are fewer productive suffixes and fewer functions (with at least one obvious exception: unlike Finnish, Estonian has productive diminutive derivation—cf. Laakso 2004). The grammaticalization of suffix chains, typical of Finnish derivation, is far less usual in Estonian, which rather shows a tendency towards restricting the number of suffixes per word. While Finnish verbal derivation often displays a concatenation of transitive and intransitive suffixes (*pieni* ‘small’ → *piene-ne*- ‘to become small(er)’ → *piene-n-tä*- ‘to make small(er)’ → *piene-n-t-y-* ‘to make oneself small(er), to become small(er)’), in Estonian such suffix combinations are far less usual. And while Estonian, similarly to Finnish and Hungarian, has a strong tradition of lexicon planning and enhancement by way of consciously created derivatives (and other neologisms), derivatives in Estonian often seem to retain a “technical” nuance: “the great majority of Estonian denominal verbs belong first and foremost to the sphere of written, official, terminological usage” (Kasik 1997: 61).

Between Finnic and Sámi the relatedness is clearly more distant than between the Finnic languages, far beyond mutual intelligibility. However, the structural similarities are clear to a bilingual speaker, and some derivational suffixes have been borrowed from Finnic to Sámi or vice versa. Structurally, Sámi derivation resembles Finnish, but especially the verbal derivation is perhaps even richer. There are a few characteristic differences such as the presence of productive deverbal inchoative derivation (North Sámi *lohka-goahhti-t* ‘to start reading’ from *lohka-t* ‘to read’; the cognate of the suffix in the same function is also present in the Finnic Veps language). For a concise but fairly detailed presentation of derivation in (North) Sámi, see Sammallahti (1998); a very detailed study, authored by a native speaker, is Ruong (1943).

26.3.2 Mordvin

Mordvin has two varieties Erzya and Moksha. The overview in this section is restricted to the Erzya dialect. Descriptive grammars from Evsejev (1934) to Grammatika (1980) treat derivational morphology in considerable detail; see also Hallap (2000). The most recent overview of Mordvin word formation can be found in Maticsák (forthcoming). The language is predominantly suffixing, the only exception being the negative prefix *a-* (stemming from a negative verbal root) as in *sati* ‘enough’—*asati* ‘not enough’.

26.3.2.1 Nominal Derivation

The most frequent Mordvin nominal suffix is *-ks*, which can be attached to nouns, adjectives, adverbs, and verbs and forms nouns with a wide range of (often lexicalized) meanings (cf. e.g. *ket’-ks* ‘bracelet’ from *ked’* ‘hand’; *pižol-ks* ‘rowan (tree)’ from *pižol* ‘rowan (berry)’; *jov-ks* ‘story, fairy-tale’ from *jovt-* ‘to tell’; *prom-ks* ‘meeting, gathering’ from *prom-* ‘to meet, to come together’). The diminutive suffix *-ka* is of Russian origin: *sazor* ‘younger sister’—*sazor-ka* ‘little sister.’ This suffix is mainly used with the names of relatives. Another suffix with similar meaning is *-iñe*: *kiska* ‘dog’—*kisk-iñe* ‘little dog.’ The two suffixes can even appear together: *numolo* ‘rabbit’—*numol-k-iñe* ‘(dear) little rabbit.’

The deverbal suffix *-ks* can be attached to intransitive verbs, yielding abstract nouns: *ozno* ‘to pray’—*oz-ks* ‘prayer’; *fizne-* ‘to grieve’—*fiz-ks* ‘sorrow.’ Action nouns are derived by means of the suffix *-mo/-ma/-me*: *mora-* ‘to sing’—*mora-mo* ‘singing.’ If attached to a stative verb the resulting derivative denotes a state: *éřa-* ‘to live’—*éřa-mo* ‘life’; *udo-* ‘sleep’—*udo-ma* ‘sleeping.’

26.3.2.2 Adjectival Derivation

The suffix *-v* is used to derive ornative or possessive adjectives from nouns. The meaning of the derivatives can be described as “relates to N” or “is made of N,” e.g. *viř* ‘wood’—*viř-ev* ‘wooden’; *kev* ‘stone’—*kev-ev* ‘stony.’ Relational adjectives are derived from nouns by means of the suffix *-ń* (identical to and possibly of the same origin as the genitive case suffix), for example *ćora* ‘man’—*ćora-ń* ‘male’; *sařar* ‘sugar’—*sařar-on* ‘sugary.’

The deadjectival suffix *-la* has an approximative function: *ečke* ‘fat’—*ečke-la* ‘fattish,’ *ožo* ‘yellow’—*ožo-la* ‘yellowish.’ Deverbal adjectives are formed by means of the suffix *-ks* which is also used to derive abstract nouns, as we saw in the previous section. Cf. *eravo-* ‘to need’—*eravi-ks* ‘needed, necessary.’

26.3.2.3 Verbal Derivation

Base verbs admit various deverbal suffixes which are used to derive Aktionsart-meanings. The suffix *-n* is a productive and frequent frequentative suffix: *korta-* ‘to talk’—*kort-ńe-* ‘to talk often,’ *čije-* ‘to run’—*čij-ńe-* ‘to run often.’ Another productive frequentative suffix is *-kšno/-kšńe* which is used to derive verbs denoting frequent but discontinuous actions: *uje-* ‘to swim’—*uj-kšńe-* ‘to swim sometimes,’ *raka-* ‘to laugh’—*raka-kšno-* ‘to laugh sometimes.’ The suffix *-še* expresses repeated prolonged actions: *pala-* ‘to kiss’—*pal-še-* ‘to kiss repeatedly,’ *mora-* ‘to sing’—*mor-še-* ‘to sing repeatedly,’ but the semantic difference, if any, between the three frequentative suffixes has not been clarified as yet.

The suffix *-ja/-je* appears in onomatopoeic verbs and loan words and carries durative meaning. The suffix *-d/-d’* can often be found in semelfactive verbs. Observe the contrast between *kučkordo-* ‘to kick with a single, sudden movement’—*kučkořa-* ‘to be kicking.’ The inchoative Aktionsart is expressed by the suffix *-kad/-gad*, for example *čara-* ‘to turn, to revolve’—*čara-kado-* ‘to start turning, revolving.’ Another frequent and productive suffix is *-t/-t’*. It creates different types of verbs depending on the phonological shape of the base. If it is attached to a stem ending in a consonant, the result is a semelfactive-perfective verb (often derived of a continuative verb, as in *korčta-* ‘to take one gulp,’ from *korča-* ‘to slurp, to drink (continuously)’), if, however, it is attached to a stem ending in a vowel the derivative has a passive-reflexive meaning, as in *mařa-* ‘to hear, to feel’—*mařato-* ‘to be heard, to be felt.’ The factitive meaning may also be expressed by the suffix *-vt/-vt’*: *keme-* ‘to believe’—*keme-vtě-* ‘to make believe.’

The suffix *-gad/-kad* can be attached to adjectives (and occasionally to nouns) to derive change-of-state verbs: *kežej* ‘angry’—*kežej-gado-* ‘to get angry,’ *koža* ‘rich’—*koža-kado-* ‘to grow rich.’ With certain adjective or numeral bases, the suffix appears as *-lgad* as in *pši* ‘hot’—*pši-lgado-*. Further suffixes with the same meaning are *-m* (*tařto* ‘old’—*tařto-mo-* ‘to grow old’), *-ija/-ja* (*čemeń* ‘rust’—*čemeń-ija-* ‘to get rusty’), *-t/-t’* (*načko* ‘damp, wet’—*nač-to-* ‘to dampen’).

26.3.2.4 Other Types of Word Formation

Back-formation, blending, and clipping are atypical processes in Mordvin but some such words are borrowed from Russian. They belong mainly to the political vocabulary.

26.3.3 Mari

There are two codified languages commonly known as Meadow (Eastern) Mari and Hill (Western) Mari. Since Meadow Mari is used by the majority of Mari speakers, in the overview we will restrict ourselves to the description of this Mari variant.

Description of word formation can be found in traditional Mari grammars, above all in Pengitov et al. (1961) and Alhoniemi (1993) and in the monograph by Lavrent'ev (1984). Riese (forthcoming) is a recent overview of word-formation in Mari. Similarly to the other Uralic languages Mari, too, has a great number of derivational suffixes; however, many of them are unproductive. Some adverbs are formed by prefixation: *ni* 'negation' (*kušto* 'where'—*ni-yušto* 'no-where'), *ala* 'some, any'—*ala-kuš* 'to where,' *keč* 'any, ever'—*keč-kunam* 'whenever,' etc. but derivation by means of prefixation is not productive.

26.3.3.1 Nominal Derivation

From among the noun forming denominal suffixes only *-aš* and *-lâk* seem to be productive. They are used to derive purposive nouns (denoting artifacts meant for N), for example *jol* 'leg, foot'—*jol-aš* 'trousers'; *šinča* 'eye'—*šinča-lâk* 'glasses.'

Deadjectival abstract nouns can be formed productively by means of the Turkic suffix *-lâk*: *jorlâ* 'poor'—*jorlâ-lâk* 'poverty,' *jočadâmâ* 'childless'—*jočadâmâ-lâk* 'childlessness.'

Deverbal nouns can be formed with the suffix *-maš*, for example *mur-aš* 'to sing' (note that *-aš* is the infinitive ending)—*murâ-maš* 'singing,' *poyân-aš* 'to gather'—*poyânâ-maš* 'gathering, meeting.' Some further deverbal suffixes are often listed as being "somewhat productive."

26.3.3.2 Adjectival Derivation

Denominal adjectives can be formed by means of various suffixes. The suffix *-an* is used to form possessive adjectives (*pâl* 'cloud'—*pâl-an* 'cloudy,' *peledâš* 'flower'—*peledâš-an* 'flowery'), the suffix *-(â)sV* is used to derive relational adjectives (*βer* 'place'—*βer-âsâ* 'local'), the suffix *-δâmV* is used to derive privative adjectives (*joča* 'child'—*joča-δâmâ* 'childless,' *püj* 'tooth'—*püj-δâmâ* 'toothless'). The suffix *-lâk* serves to derive purposive adjectives ('for the use of sg,' e.g. *kornâ* 'journey'—*kornâ-lâk* 'for the journey').

26.3.3.3 Verbal Derivation

From among the denominal verbal suffixes *-l* is the most common and productive, for example *šük* 'litter'—*šük-l-aš* 'to drop litter,' *târma* 'harrow'—*târma-l-aš* 'to harrow.' Change of state verbs are derived from adjectives by means of the suffix *-em*, for example *ošâ* 'white'—*oš-em-aš* 'to become white, or by means of the suffix *-aŋ*, for example *lopka* 'wide'—*lopka-aŋ-aš* 'to become wider.' As in other Uralic languages deverbal derivation is by far the richest in possibilities. Frequentative verbs can be derived by means of the suffix *-âl* as in *âšt-aš* 'to do'—*âšt-âl-aš* 'to do it often,' semelfactive verbs by means of the suffix *-al* as in *onč-aš* 'to look'—*onč-al-aš* 'to give a look,' medial verbs by means of the suffix *-alt*, as in *âšt-aš* 'to do'—*âšt-alt-aš* 'to be done,' causative verbs by means of the suffix *-âkt* as in *âr-aš* 'to become warm'—*âr-âkt-aš* 'to warm up.'

26.3.3.4 Adverbial Derivation

Adverbs can be formed from nouns with certain case suffixes, for example (i) with the genitive case suffix *-n/-ân* (*jol* 'foot'—*jol-ân* 'on foot'), (ii) with the accusative case suffix

-*m/-â**m* (*telâ* ‘winter’—*tel-â**m* ‘in winter’), (iii) with the inessive case suffix -*šte/-što/-štö* (pronunciation of *all* is -*štâ*) (*šola* ‘left’—*šola-štâ* ‘on the left’), etc.

26.3.3.5 Conversion

As in other Finno-Ugric languages, nouns and adjectives are formally not distinguished, the category “nomen” is indeterminate between noun and adjective. This means that adjectives, too, can take nominal suffixes. Otherwise, conversion can operate in various directions: participle to noun (e.g. *tunâkt-aš* ‘to teach’—*tunâktâ-šâ* ‘teaching (one)’ → ‘teacher’), verb to noun (e.g. *poγ-aš* ‘to collect, to accumulate’—*poγâ* ‘property’ → ‘treasure’), nouns to verb (e.g. *lum* ‘snow’—*lum-aš* ‘to snow’; this, however, is lexically restricted to a few ambiguous N/V stems), adjective to verb (e.g. *šapalyâ* ‘pale’—*šapaly-aš* ‘to turn pale’). Sometimes also verb to verb conversion (verbs shifting between the two conjugations) is considered to be a genuine case of conversion.

26.3.3.6 Reduplication

Total reduplication of words is sometimes (but not systematically) used in Mari to express intensity.

26.3.4 Permic

The Permic branch of Uralic consists of Udmurt (also known as Votyak) and Komi (also known as Zyryan, standardised in two varieties: Komi-Zyryan (in the Komi republic) and Komi-Permyak. Information on Permic word formation can be found in teaching materials (Fedyuneva 2000, Fedjuneva and Punegova 2002, Karmanova 2002). For an overview of Permic word-formation, cf. Bartens 2000: 125–9, 139–41, 273–91.

Though the Permic languages have a similar inflectional system, the derivational suffixes differ considerably. Komi has borrowed derivational suffixes from Russian, Udmurt from Tatar.

26.3.4.1 Nominal Derivation

The Russian influence on Komi can clearly be seen in the cases where no corresponding suffix can be found in Udmurt. Such is the case with the suffix -*ñik* (for males) and -*ñitsa* (for females) as in *porš-ñik* ‘swineherd’—*porš-ñica* ‘female swineherd’ (from *porš* ‘pig’) or with the diminutive suffix -*ka* as in *njv* ‘girl’—*njv-ka* ‘girlie.’ As expected, nouns derived from nouns can express a great variety of meanings.

Both Permic languages can form abstract nouns from adjectives. In Komi such suffixes are -*a* (*kjz* ‘thick’—*kjz-a* ‘thickness’) and -*ed* (*kokñi* ‘light’—*kokñed* ‘lightness’), both of which also appear in adjective derivation; in Udmurt we find -*dala* (*kuz* ‘long’—*kuz-dala* ‘length’) and -*lik* (as in Mari, borrowed from Turkic: *zem* ‘true’—*zem-lik* ‘truth’). Especially in Komi, some nouns seem to be in the process of being grammaticalized into noun suffixes, for example *tor* ‘piece’: *jëv-tor* ‘some milk, a drop of milk.’

Action nouns are derived from the verbal base by means of the suffix *-(o)n* in Udmurt and by means of the suffix *-em* in Komi: *vu-* ‘to arrive’—*vu-on* ‘arrival’ (Udmurt) and *puk-* ‘to sit’—*puk-em* ‘sitting’ (Komi). Komi *-em* also functions as the suffix of the past participle, while Udmurt *-on*, although labeled as a “participle” suffix in some grammars, should rather be classified as an “action nominal” suffix (Bartens 2000: 242).

Agent nouns are formed with Udmurt *-(i)ś*, Komi *-jš* (as e.g. in Hungarian, the suffix is formally identical to the present participle suffix): *djšet-* ‘to teach’—*djšet-iś* ‘teacher’ (Udmurt) and *veļed-* ‘to teach’—*veļed-jś* ‘teacher’ (Komi).

26.3.4.2 Adjectival Derivation

Denominal adjectives can express (i) a possessive meaning (sg has the property X), (ii) a privative meaning (sg lacks the property X) or (iii) are relational (mostly temporal or local). Possessive adjectives are formed with *-o* (Udmurt) / *-a* (Komi), e.g. *šundi* (Udmurt) / *šondi* (Komi) ‘sun’—*šundj-jo* / *šondi-a* ‘sunny’. The privative suffix is *-tem* in Udmurt and *-tem* in Komi as in *dun* (Udmurt) / *don* (Komi) ‘price’—*dun-tem* / *don-tem* ‘cheap’. Examples for relational suffixes in Komi are *-ša* (temporal) as in *talun-ša* ‘of the current day’ (*talun* ‘today’) and *-sa* (local) as in *śikt-sa* ‘from the village, in the village’ (*śikt* ‘village’). Adjectives can also be formed from base adjectives with various suffixes; deadjectival adjectives express approximation, diminution, or endearment.

26.3.4.3 Verbal Derivation

Both Udmurt and Komi have a host of suffixes which derive verbs from diverse bases. It is not unusual that the same suffix can be attached to both nominal and verbal stems. For example, Udmurt *-a(l)-*, Komi *-al-* ~ *-av-*³ can form causative or instrumentative denominal verbs (Udmurt *ńima-* ‘to name’ from *ńim* ‘name’, Komi *kemav-* ‘to put on shoes’ from *kem* ‘shoe’), translative (“inchoative”) verbs from adjectives (Udmurt *ćebera-* ‘to become beautiful’ from *ćeber* ‘beautiful’), and frequentative verbs from verb stems (Komi *usal-* ‘to fall often’ from *uś-* ‘to fall’). The meanings of the derivatives are not always predictable, and there is not sufficient data available about the productivity of individual suffixes.

The functions of the deverbal verb suffixes are difficult to describe, as they often express sophisticated combinations of aspect and Aktionsart. Frequentative suffixes may be used to express the plurality of object (Udmurt *l’ukilj-* ‘to divide in many parts’, from *l’uk-* ‘to divide’) or subject (Komi *voav-* ‘to come (of many), to keep arriving’, from *vo-* ‘to come’); this comes close to syntactic number agreement and might, thus, be considered part of the inflectional rather than derivational morphology. However, the frequentative *l* in Komi can also express restricted duration or place the event in distant past: A. S. Puškin *čuž-l-is* 1799 *vojn* ‘A. S. Pushkin was born in the year 1799’ (SKJa 1955: 224). Both Komi and Udmurt have suffixes for deriving momentaneous/diminutive and/or

³ In (Standard) Komi, non-syllable-final *l* regularly alternates with syllable-final *v*.

inchoative deverbal verbs (e.g. Udmurt *lobd'žj-* / Komi *lebžj-* 'to set off flying' from *lob-/leb-* 'to fly,' Komi *vorsjšt-* 'to play (a little, quickly)' from *vors-* 'to play').

Both Permic languages can also derive reflexive (Komi *ošji-š-* / Udmurt *ušja-ški-* 'to boast, to praise oneself' from *oški-/ušja-* 'to praise') and curative-causative (factitive) verbs (e.g. Komi *berd-əd-* / Udmurt *berdj-tj-* 'to make (somebody) cry' from *berd(i)-* 'to cry').

26.3.4.4 Other Types of Word Formation

Conversion is not a typical means of word formation in Permic. Back-formation, blending, and clipping may occur occasionally but are certainly atypical of Permic morphology.

Reduplication of whole words indicates intensity (Udmurt *zol* 'strong'—*zol-zol* 'very strong' and Komi *il'in* 'far'—*il'in-il'in* 'very far').

26.3.5 Hungarian

Detailed description of Hungarian word formation can be found in various descriptive grammars. The most detailed discussion of word-formation suffixes—including obsolete and unproductive ones—can be found in Tompa (1961). More recent works are Keszler (2000) and É.Kiss et al. (2003). An overview of productive word-formation processes is provided in Kiefer (forthcoming).

26.3.5.1 Noun Derivation

Nouns can be derived from nouns, adjectives, and verbs. In Hungarian frequent and productive denominal derivational suffixes are (i) *-ka* or *-(V)cska* for deriving diminutives (e.g. *asztal-ka* 'small table,' *kert-ecske* 'small garden'), the distribution of the two forms is largely determined by phonological factors, (ii) *-ság* to derive nouns denoting types of trades (e.g. *asztalos-ság* 'joiner's trade') and (iii) *-(V)s* to derive names of profession (e.g. *posta-s* 'postman').

The deadjectival suffix *-ság* is used to derive nouns denoting abstract properties: *szép-ség* 'beauty,' *lusta-ság* 'laziness.'

The deverbal action nouns are derived by means of the suffix *-ás* (e.g. *olvas-ás* 'reading') and deverbal agent nouns by means of the suffix *-ó* (e.g. *olvas-ó* 'reader'). Action nouns may denote a single event, a complex event, or the result of an activity. Nouns derived from particle verbs normally have a complex event reading. (E.g. *ír-ás* 'writing' simple event or result, *meg-ír-ás* 'writing in full' complex event reading.) The suffix *-ó* is used to derive nouns from verbs with the following meaning: (i) agent nouns such as *tanít-ó* 'teacher'; (ii) instrument nouns such as *vés-ő* 'chisel'; (iii) location *vár-ó* 'waiting room'; (iv) simple events such as *válogat-ó* 'selection match.'

26.3.5.2 Adjective Derivation

Adjectives can productively be derived from nouns and verbs. From among the denominal adjectivalizers the suffix *-i* is mainly used to derive relational adjectives. The

following types of nouns may serve as the base for the derivation: names of professions (*professzor-i* ‘professorial’), names of institutions (*egyetem-i* ‘pertaining to the university’), geographical objects (*párizs-i* ‘Parisian,’ *duna-i* ‘Danubian’), arts and sciences (*építészet-i* ‘architectural,’ *történelm-i* ‘historical’), temporal periods or points (*reggel-i* ‘pertaining to morning,’ and family names (*shakespeare-i* ‘Shakespearean’). Syntactically, the main characteristics of relational adjectives has to do with their non-predicative use and their non-scalarity. Semantically, the opposition descriptive/relational adjective takes into account the referential properties of the adjective. While descriptive adjectives modify the referent, relational adjectives modify the reference of the nominal head. Compare *nehéz feladat* ‘difficult task, the task is difficult’) and *professzori feladat* ‘professorial task, *the task is professorial’). Excluded as basis are, among others, mass nouns (**arany-i* ‘golden’) and names of animals and plants (**macska-i* ‘feline,’ **rózsa-i* ‘of roses’).

The suffix *-(V)tlAn* is used to derive privatives from concrete nouns (*csillag-talan* ‘starless’) and antonyms from derived adjectives (*történeti-etlen* ‘unhistorical’). It can also be attached to base verbs: *olvas-atlan* ‘unread,’ where the derived adjective is the antonym of the participial adjective *olvas-ott* ‘read,’ that is, *olvasott—olvasatlan* form antonymous pairs.

Adjectives can be derived from present participles if the verb has a Source subject argument and an Experiencer argument. We thus get adjectives such as *dühít-ő* ‘enraging,’ *fáraszt-ó* ‘tiring.’ This case can also be interpreted as conversion from present participles to adjectives.

The suffix *-hAtÓ* is a composite suffix composed of the verbal root *-hAt* ‘may, can’ and the participial suffix *-Ó*, the meaning of the composite suffix can be rendered by the English suffix *-able*. This suffix can be attached to transitive verbs: *olvas-ható* ‘readable,’ *elemez-hető* ‘analyzable.’

The suffix *-(j)Ű* is attached to phrases rather than to words and it is used to derive prenominal attributes: [*hosszú haj*]-*ú* ‘long haired,’ that is, the suffix *-(j)Ű* is strikingly similar to the German suffix *-ig* in *langhaarig* and English *-ed* as in *long haired*.

26.3.5.3 Verbal Derivation

Hungarian has a very rich system of verbal derivations, it has more than a dozen productive verbal derivational suffixes. Among the most frequent suffixes we find the denominational suffixes *-(V)z* (*golf-oz(ik)* ‘play golf’) and *-(V)l* (*sport-ol* ‘do sports’). The distribution of the two suffixes is largely determined by the phonological shape of the base.

Deadjectival verbs are derived from monosyllabic stems by means of the suffix *-Ul* (*szép-ül* ‘to become prettier’) and from polysyllabic stems by means of the suffix *-Vdik* (*magas-odik* ‘to become higher’). The suffix *-ít* is the causative equivalent of these suffixes (*szép-ít* ‘make prettier,’ *magas-ít* ‘make higher’).

The deverbal productive suffixes include the attenuative/frequentative suffix *-(V)gAt* (*olvas-gat* ‘to read leisurely,’ *énekel-get* ‘to sing a little bit’), the causative suffix *-(t)At* (*olvas-tat* ‘to let read,’ *énekel-tet* ‘to make sing’) and the obsolete passive suffix *-(t)Atik* (*olvas-tatik* ‘to be read’). The suffix *-ál* is used to accommodate foreign roots into the

Hungarian verbal system: *install-ál* ‘to install,’ *edit-ál* ‘to edit.’ A verbal root can take up to four derivational suffixes: the causative, the factitive, the attenuative/frequentative, and the possibility suffix (in that order).

Particle verbs are not words but rather constructions (Booij 2010): particles can be separated from their base under various syntactic conditions. However, particles can be reduplicated to express iterativity (the iterative Aktionsart) and the reduplicated particle can no longer be separated from the base. Sentence (1a), if negated, becomes (1b) where the particle is postposed. Sentence (1c) contains the reduplicated particle but this sentence cannot be negated (by internal negation) since the negative particle, too, must occupy the preverbal position which is no longer “free” since it is already taken by the reduplicated particle. Note that the reduplicated particle can no longer be postposed, (1e) is ungrammatical.

- (1) a. *El-olvassa a könyvet.*
 Part-read the book-acc
 ‘He/she will read the book through’
- b. *Nem olvassa el a könyvet.*
 Neg read Part the book-acc
 ‘He/she will not read the book’
- c. *El-el-olvas egy könyvet.*
 Part-Part read a book-acc
 ‘He/she reads a book (from time to time)’
- d. **Nem el-el-olvas egy könyvet.*
 Neg Part-Part read a book-acc
 Lit. ‘He/she does not read a book (from time to time)’
- e. **Olvass el-el egy könyvet.*
 Read Part-Part a book-acc

26.3.5.4 Adverbial Derivation

The most frequent productive adverbial suffixes are *-(A)n* (*meleg-en* ‘warm’ Adv), *-Ul* (*lengyel-ül* ‘in Polish’ Adv) and *-lAg* (*elmélet-i-leg* ‘theoretically’). Adverbs and adjectives are systematically distinguished morphologically. The distribution of the various forms is fully predictable. The suffix *-(A)n* is attached to base adjectives as well as to certain derived adjectives. Negative adjectives as well as ethnic adjectives take the suffix *-Ul*, and adjectives ending in the derivational suffix *-i* or *-beli* take the adverbial suffix *-lAg*.

26.3.5.5 Conversion

From the point of view of inflectional morphology nouns, adjectives and pronouns behave similarly. That is why Hungarian grammars subsume them under the label

“nominal catagories.” This explains also why Hungarian grammars do not speak of conversion in the case of base nouns and adjectives. Words denoting nationality such as *magyar* ‘Hungarian,’ *német* ‘German’ can equally function as nouns and adjectives. This is equally valid for color words: *piros* ‘red,’ *kék* ‘blue.’ Yet there are some cases where genuine conversion occurs. Adjectives can be derived from nouns by means of the suffix *-i* with a number of different meanings but the derived adjective can be converted into a noun in the case of the meaning “coming/stemming from N”: *város* ‘town’—*város-i* ADJ ‘coming/stemming from a town’—*város-i* N ‘person coming from or living in a town,’ but if the noun denotes a profession or occupation the adjectival form can never be converted to a noun: *professzor*—*professzor-i* ADJ—**professzor-i* N. The adjectival suffix *-(V)s* can be used to derive adjectives with the meaning “being associated with N” and the adjective obtained can be then turned into a noun: *újság* ‘newspaper’—*újság-os* ADJ ‘having to do with newspapers’ (as in *újságos-bódé* ‘newspaper-stall’)—*újság-os* N ‘newsagent.’ But once again not all adjectives derived with the suffix *-(V)s* can be converted to nouns in this way: *láz* ‘fever’—*láz-as* ADJ ‘feverish’—**láza-as* N.

26.3.5.6 Other Types of Word Formation

Back-formation and blending do occur but only sporadically. An example of back-formation is *tény-feltárás* lit. ‘fact finding’ from which the verb *tényfeltár* ‘establish the facts’ can be derived. Occasionally we may also encounter blending: *film* + *internet* → *filinternet* ‘internet café in a movie theatre’ but such formations are rare and do not get lexicalized (Ladányi (2007)). Neither back-formation nor blending is typical of Hungarian derivational morphology. Clipping, however, is a semi-productive process and it follows the pattern of hypocoristic formation: the diminutive suffix *-i* is attached to the maximal initial syllable: *doktor* → *dok+i*, *paradicsom* ‘tomato’ → *par+i*, *sütemény* ‘cake’ → *süt+i*.⁴

26.3.6 Nenets

Nenets is divided into two major dialects: Tundra Nenets and Forest Nenets. From these two dialects Tundra Nenets has been more fully described. The overview in this section concerns Tundra Nenets. Derivation is the most typical method of enriching the vocabulary in Nenets. Compounding is rare and it is mostly restricted to Russian loanwords.

Earlier works on Nenets describe the inflectional system of the language but not word formation. Some information, however, can be found in textbooks (e.g. Kuprianova, Barmič, and Homič 1985). Wagner-Nagy (forthcoming) provides a useful overview of Nenets word-formation.

26.3.6.1 Noun Derivation

Among the denominal noun-forming suffixes, the augmentative and the diminutive suffixes are the most productive. The augmentative suffix is *-ʔjA*, for example *maʔ*

⁴ The “maximal initial syllable” is defined as the first syllable of a word plus the following consonant.

'tent'—*maʔja* 'big tent,' the diminutive suffix is *-ko*, for example *ńawa* 'hare'—*ńawa-ko* 'small hare.'

Action nouns are derived from the verbal base by means of the suffix *-wa*, for example *jiʔe-sʹ* 'to live'—*jiʔe-wa* 'life.' Agent nouns are derived from the nominalized form of the present participle ending in *-na*, *-da*, or *-ta*. For example, *xańe-sʹ* 'to hunt'—*xańe-na* 'hunter.' Contrary to the closely related Nganasan, Nenets does not differentiate morphologically between the occasional performer of an action and names of professions. The instrument of an action can be derived by means of the suffix *-cʹ*? or *-sʹ*? as in *tir-cʹ* 'to comb'—*tir-cʹ* 'brush.'

26.3.6.2. Adjectival Derivation

Base adjectives are rare; adjectives are frequently formed from nouns. The suffixes *-ki/-gi* and *-i*, *-ij*, *-ńij*, *-ńgi* are used to derive relational adjectives with the meaning "belonging to a certain place," e.g. *mʹa?* "tent" - *mʹa-ki* "belonging to the tent." Nenets has a similitive suffix *-raxa/-laxa/-rx* expressing the meaning "being like N," e.g. *mʹa?* "tent" - *mʹa?-laxa* "similar to a tent, tent-like." The suffix *-sawej* may be used to derive possessive adjectives, e.g. *ńe* "woman"—*ńe-sawej* "with his wife, married." The privative suffix is *-sʹa(da)*, e.g. *ńe-sʹa* "without woman, unmarried."

The derivation of deadjectival and deverbal adjectives does not seem to be productive.

26.3.6.3 Verbal Derivation

Nenets is rich in productive verb-forming suffixes which are either denominal or deverbal, the derivation of deadjectival verbs seems to be unproductive. Throughout Uralic the same three suffixes *-ta/-da*, *-li* and *J* occur with very diverse semantics. They can also be found in borrowed verbs (Laakso 1997). But it is unclear to what extent they are productive. Of the suffixes with a more clear-cut semantics the suffix *-m?* is used to derive translative verbs ("to become more X"). This suffix also attaches to the plural genitive stem, as in *sera* 'widow'—*serʹ em* 'to become a widow,' *pir* 'height'—*piram* 'to become tall.' A typical Samoyedic derivation is to form derivatives with the meaning "it smells like X." The Nenets suffix, for example, used to this end is *-ja*, *xalʹa* 'fish'—*xalʹa-ja* 'to smell like fish,' *xerpsin* 'kerosene'—*xerosin-ja* 'to smell of kerosene.' The derivation seems to be productive in the domain of typical smells.

Nenets has a considerable number of productive deverbal suffixes. The transitive-causative *-ta/-da* is used to form transitive verbs, for example *xawa* 'to fall'—*xawa-da* 'to drop,' *xaa* 'to die'—*xaa-da* 'to kill.' In addition to *-ta/-da*, the suffixes *-tʹe/-dʹe*, *-laa* and *-lʹe*, too, can be used to derive transitive verbs from intransitives. However, the distribution and the degree of productivity of the transitivizing suffixes remains unclear.

As in all Northern Samoyedic languages, most Nenets verbs are perfective and the corresponding imperfective verbs are formed by various derivational suffixes, which normally carry an Aktionsart meaning as well. Productive and frequent suffixes are the suffix *-rʹ?* used to derive frequentative-iterative verbs (e.g. *ńam* 'to eat up'—*ńaw-orʹ* 'to be eating'), the suffix *-ńka* used to derive durative-progressive verbs (e.g. *to* 'to arrive'—*to-ńga* 'to be approaching'), the habitative suffix *-sʹeti/-cʹeti* (e.g. *nu-* 'to stand'—*nu-sʹ*

iti ‘to use to stand’), the diminutive (attenuative) suffixes *-jbt’è/.jebt’è/-bt’è* (e.g. *je* ‘to be ill’—*je-jebt’è* ‘to be a bit ill’).

There are two instances where verbs take prefixes, namely a form of the negative auxiliary *ñiis’* which can further be prefixed by the emphatic prefix *wu-* or the interrogative prefix *xa-* yielding complex prefixes, however complex prefix formation cannot be considered to be productive since such prefixes can only be attached to the negative auxiliary.

26.3.6.4 *Adverbial Derivation*

Adverbs of manner can be formed from an adjective by means of the suffix *-wna*, for example *ɲarka* ‘big’—*ɲarka-wna* ‘to a great extent,’ *sawa* ‘good’—*sawa-wna* ‘well.’ This suffix is formally identical with the prolative inflectional suffix, which shows, once again, that the border between adverbial and case suffixes is fuzzy.

26.3.6.5 *Conversion*

As in the other Uralic languages, noun to adjective conversion is very common in Nenets, the class of adjectives is often considered to be a secondary word class. In the construction “nominal attribute + noun” the attributive noun is normally interpreted as a qualitative adjective: for example *jal’a* ‘sun’—*jal’a m’akana* ‘bright tent.’ Under certain conditions present participles of stative verbs can be converted to adjectives: for example *saxarja-s’* ‘to be sweet’—*saxarja-da* ‘being sweet’ (participle) → ‘sweet’ ADJ.

26.4 SOME CONCLUSIONS

From the survey of some productive derivational processes in various Uralic languages the following picture emerges.

1. Derivation in Uralic is predominantly suffixing. There are a few exceptions only, which represent a special type of derivation such as negative prefixing.
2. Multifunctional suffixes are not infrequent. From among these suffixes, those attachable to nouns and adjectives are more frequent than ones which can be attached to all three major categories.
3. Deverbal suffixes are more varied than denominal suffixes and denominal suffixes are richer than deadjectival suffixes.
4. Throughout Uralic, certain functions of derivation are well represented, sometimes expressed with ancient, etymologically cognate suffixes:
 - deverbal agent and action nouns;
 - deverbal verbs: almost all Uralic languages can derive factitive (curative-causative) and reflexive, passive or medial verbs; verbal derivation is also used to express diverse forms of aspect and Aktionsart (frequentative/iterative, momentaneous/semelfactive, etc.);

- denominal or deadjectival verbs for causing (“causative”) or becoming something (translative/inchoative);
 - deadjectival (abstract) nouns and denominal (possessive, relational) adjectives; all Uralic languages also form privative (caritive) adjectives;
 - diminutive noun derivation occurs in most (although not all) Uralic languages, while productive augmentative derivation is only known in Samoyedic;
 - adverbs are distinct from adjectives, and typically there are distinct suffixes for adverb derivation; the boundary between these adverb suffixes and local or adverbial case suffixes may be contested or fuzzy.
5. The status of conversion in Uralic derivation is debatable. The boundary between nouns and adjectives is not clear-cut, present participles are often functionally or even formally identical with agent nouns, and other participles as well can be lexicalized into adjectives, while derived adjectives can be lexicalized to noun-like use. In most Uralic languages there is a small, closed, and lexically restricted group of underived word stems that can be inflected either as nouns or as verbs, and sometimes, due to phonological reasons, verbal and nominal derivation suffixes coincide, giving rise to identical derived verb and noun stems. Whether these cases are regarded as conversion is a matter of analysis.
 6. Reduplication is not typical of the whole language family. In Hungarian verbs particle reduplication is a productive process, the derivatives express iterated action. In Permic (and Mari) whole words can be repeated to express intensity (Stolz et al. [2011] consider this an areal phenomenon, also typical of the neighboring Turkic languages). Neither the Hungarian verbs with reduplicated particle nor the Permic iterated words can be lexicalized.
 7. Back-formation, blending, and clipping are not characteristic features of Uralic.

ACKNOWLEDGMENTS

We wish to express our gratitude to László Fejes for his contribution to part of the chapter on Permic, to Timothy Riese for his contribution to part of the chapter on Mari, to Sándor Maticsák for his contribution to part of the chapter on Mordvinic, and to Beáta Wagner-Nagy for her contribution to part of the chapter on Nenets.

CHAPTER 27

ALTAIC

IRINA NIKOLAEVA

27.1 ALTAIC LANGUAGES

ALTAIC is a so-called macro-family, that is, a grouping of several traditionally defined linguistic families that are said to descend from a common ancestor, the hypothetical Proto-Altaic. It is highly controversial as a genetic entity: there is no consensus on whether the genetic relationship between Altaic languages is proven and, if so, which linguistic families the grouping may include.

In its “classic” form, the Altaic macro-family comprises three well-established genealogical units, the Turkic, Mongolic, and Tungusic families of languages. This grouping is sometimes referred to as “Micro-Altaic.” It was explored in the classical works of G. J. Ramstedt and N. Poppe in the mid 20th century (Ramstedt 1952, 1957, 1966, Poppe 1960, and other works) and remained widely popular for some time although, as Georg et al. (1999: 73–4) emphasize, it has almost no supporters at present. The expanded version of the Altaic theory, which may be referred to as “Macro-Altaic,” additionally includes Korean and Japanese. This view is probably the most widespread today and is represented by the works of S. Starostin (1986, 1991, 2000, 2005), Blažek (2006), Robbeets (2004, 2005, 2007), A. Dybo (1996, 2000, and other work), and Dybo and Starostin (2008), among others. The substantial volume *An Etymological Dictionary of the Altaic Languages* (Starostin et al. 2003) presents 2,800 cognate sets for the Macro-Altaic languages, as well as an elaborate system of sound correspondences and some grammatical parallels. An alternative classification groups together the Micro-Altaic languages, on the one hand, and Korean, Japanese, and Ainu, on the other hand, linking these two groupings within an even larger genealogical entity. This view is represented by Street (1962), Patrie (1982) and partly by Greenberg (2000–2), but is not widely supported.

The opponents of the Altaic theory in any of its forms accept that Altaic languages exhibit a remarkable number of lexical and morphological cognates and that it is

generally possible to establish more or less regular sound correspondences based on these cognates. The question is how to judge these commonalities: as genetic inheritance (the result of the divergence of a common proto-tongue) or as areal convergence leading to extensive borrowing. According to the anti-Altaic view represented by Clauson (1959, 1968, and other work), Doerfer (1963, 1974), and, more recently, Georg (1999/2000, 2004, 2005) and Vovin (2005), similarities between Altaic and especially Micro-Altaic languages are due to relatively recent contacts, which are often limited to certain languages within the family, and may also reflect chance resemblances.

This chapter deals with Micro-Altaic, that is, Turkic, Mongolic, and Tungusic languages, which show a similar typological profile and may indeed represent a kind of large *Sprachbund* (whether or not they are genetically related). They will be referred to simply as “Altaic” in this chapter. The families comprise about 66 living languages spoken in northeast Asia through Central Asia to Anatolia and Eastern Europe, and have a total of about 348 million speakers at the present time.

The Turkic family is composed of a large number of fairly closely related languages. The traditional classification (e.g. Johanson and Csátó 1998: 82) distinguishes between the following branches of this family, with an indication of the major languages of each branch:

- (1) Southwestern or Oghuz group
 - Azerbaijani, Crimean Tatar, Salar, Turkish, Turkmen
- Southeastern or Uyghur group
 - Chagatay (extinct), Uyghur, Uzbek
- Northwestern or Kipchak group
 - Kuman, Karaim, Kumyk, Karachai-Balkar, Tatar (with several varieties), Bashkir, Karakalpak, Kazakh, Kyrgyz
- Northeastern or Siberian group
 - Altai Oirot (with several varieties), Khakas, Shor, Tofa, Tuva, Sakha/Yakut, Chulym, Dolgan
- Bulghar group
 - Chuvash, possibly a number of extinct languages
- Arghu group
 - Khalai

The Ethnologue classification (Lewis 2009), based on Baskakov (1962) and Deny et al. (1959–64), refers to the Southwestern group as Southern, to the Southeastern group as Eastern, to the Northwestern group as Western, and to the Northeastern group as Northern. Khalai is included in the Southern group.

There are about a dozen extant Mongolic languages. The classificational scheme in (2) is based on Lewis (2009):

(2) Eastern group

Dagur/Daur

Monguor

Bonan, Dongxiang, Kangjia, Tu, Yugur

Oirat-Khalkha

Buryat, Mongolian proper, Darkhat, Kalmyk-Oirat

Western group

Mogholi/Moghol

Janhunen (2006) refers to the Oirat-Khalkha sub-group as Central Mongolic and additionally lists Khamnigan and Ordos as belonging to this group, while in the Ethnologue classification they are treated as varieties of Mongolian proper. The Monguor sub-group is referred to as Shirongolic and is said to include Monguor/ Mongghul/ Mangghuer as the fifth language. In fact, Mongolic languages demonstrate a relatively close mutual relationship, so many of their features are better described in terms of some kind of wave model of linguistic geography than a tree model (Janhunen 2003: xvii). The known extinct Mongolic languages are Middle Mongolian (spoken from the 13th until the early 15th or late 16th century) and Classical Mongolian (from the 17th century until the 20th century).

The scheme in (3) represents the traditional classification of the Tungusic languages based on Cincius (1949), Benzing (1955: 9–10), Menges (1978: 370), Comrie (1981: 58), and Lewis (2009).

(3) Northern group

Evenki, Even, Solon, Neghidal

Southern group

Southeastern

Ulch, Nanai, Oroch, Oroch, Udeghe/Udihe

Southwestern

Manchu, Jurchen, Xibe

The alternative classifications treat Manchu as an independent branch (Vasilevič 1960, Avrorin 1959–61, Sunik 1962: 16–20, 1968: 54, 1997: 154). On some other classifications of Tungusic and/or comparisons between them see Ikegami (1974), Doerfer (1978), Vovin (1993), and Miller (1994).

The chapter is based on the following sources. Unless indicated otherwise, the Turkic data come from Turkish (Kornfilt 1997, Göksel and Kerslake 2005), Chatagay (Bodrogligeti 2001), Modern Uyghur (Jong 2007), Kazakh (Balakaev 1962), Sakha (Ubrjatova 1982), and Salar (Tenishev 1976), that is, most branches of the family are represented. General information on Mongolic languages is provided after Jarceva (1997) and Janhunen (2003), and I also cite data from Mongghul (Slater 2003), Dagur (Martin 1961), Buryat and Kalmyk (Charekov 1990). The Tungusic materials are cited after Avrorin (1959, 1981) for Nanai, Nikolaeva and Tolskaya (2001) for Udihe, Nedjalkov (1994, 1995) and Boldyrev (1976, 1987) for Evenki, Malchukov (1995), Novikova et al. (1991), Cincius (1947), Benzing (1955) for Even, and Avrorin (2000) for Written Manchu.

When I cite examples from the sources which employ the Latin-based transcription, the transcription of the source remains unaltered. Examples from the Cyrillic-based sources are transliterated using the traditional Altaic transcription. Capital letters in the abstract phonological representations denote harmonizing segments, for example *A* stands for a non-high harmonizing vowel, while *U* and *I* denote high harmonizing vowels.

27.2 WORD FORMATION IN ALTAIC

In this section I provide a general overview of the most important derivational processes in Altaic that affect the major grammatical classes of nouns, adjectives, and verbs. As mentioned in Section 27.1, Altaic languages share numerous typological similarities. They are consistently SOV, whereas verb finality correlates with the head-finality evident in noun phrases, postpositional phrases and nonfinite dependent clauses. The morphological structure is characterized by a rather high degree of agglutination. In most cases the words can be easily divided into a linear sequence of distinct morphemes, each of which has a fairly consistent shape and a single function. The only native type of affixes are suffixes, that is, most bound morphemes follow the root, although there are some rare instances of prefixation. Other derivational processes are compounding, conversion, and (rarely) reduplication. Non-concatenative morphology is overall not typical.

27.2.1 Affixation

Most derivational morphology of modern Turkic, Mongolic, and Tungusic languages is quite old, that is, inherited from the respective proto-languages. The languages are rich in interclass (category-changing) derivational processes, whereby affixation marks the transition of a lexical item into another grammatical class. Historically, Altaic languages were also rich in category-preserving derivational processes, but relatively few of them are productive in the modern languages. Many suffixes from this latter group are etymologically complex and go back to a combination of two (or more) old derivational morphemes.

27.2.1.1 *Category-changing Suffixation*

The three open word classes (nouns, verbs, and adjectives) can be distinguished by morphological and syntactic criteria. Nouns are characterized by grammatical categories of number (unmarked singular and marked plural), case, and possessive. Possessive affixes (or clitics, in some languages) attach to the head of a possessive NP and cross-reference the possessor for person and number. The case system includes between six and thirteen to fourteen grammatical cases, depending on the language. Gender and definiteness are not grammatical categories in Altaic.

Nouns can be derived from verbs, adjectives, and (less frequently) from adverbs. Deverbal nouns are formed by derivational processes with restricted productivity and show a tendency for lexicalization. They have all the properties of non-derived nouns and differ in this from nonfinite verbal forms, which are also derived from verbs by various nominalizing processes but exhibit a mixture of nominal and verbal properties: for instance, they typically head clauses, take verbal morphology such as the passive, and preserve the argument structure of the base verb. “True” deverbal nouns denote the process or result of X, where X is the verbal base (*nomina acti*) (e.g. Uyghur *bas-* ‘print’ > *bas-ma* ‘printing,’ Chagatay *bil-* ‘know’ > *bil-im* ‘knowledge,’ Written Manchu *efi-* ‘play’ > *efi-n* ‘game’), the performer of X (*nomina agentis*) (Uyghur *yaz-* ‘write’ > *yaz-ghuchi* ‘writer,’ Even *xupkut-* ‘teach’ > *xupkuč-imŋe* ‘teacher’), the place where X occurs or the instrument (Turkish *bat-* ‘sink’ > *bat-ak* ‘marsh, swamp,’ Even *muli-* ‘fetch water’ > *muli-mkin* ‘the place where one takes water’), an entity with inherent (and often negative) quality related to X (Turkish *unut-* ‘forget’ > *unut-kan* ‘forgetful person,’ Even *deb-* ‘eat’ > *debu-tur* ‘glutton’), as well as the result or object of X (*nomina patientis*) (Kazakh *tanu-* ‘know’ > *tan-ıs* ‘acquaintance,’ Even *tore-* ‘talk’ > *tore-n* ‘word, speech’). The relevant derivational suffixes are often polyfunctional. Nouns derived from adjectives are less frequent, but do occur, for example Uyghur *batur* ‘brave’ > *batur-luq* ‘bravery,’ Kazakh *aq* ‘white’ > *aq-tıq* ‘whiteness,’ Turkish *iyi* ‘good’ > *iyi-lik* ‘goodness’.

As a grammatical class, Altaic adjectives cannot generally be distinguished from nouns based on inflectional criteria alone because they have similar inflectional properties to nouns, in the sense that they take the same case, number, and possessive morphology when functioning within a phrase without a semantic head. However, adjectives clearly differ in terms of their syntactic distribution. In particular, their main function is attributive modification and in this function they do not impose any additional marking of syntactic dependency. The attributive concord in case and number is only present in Northern Tungusic, but completely unknown in other Altaic languages. Unlike nouns, adjectives can be used in the copredicative function. Manner adverbs modify verbs or adjectives but never nouns. Altaic adjectives fall into two large morphological classes: underived and derived adjectives. Underived adjectives are morphologically simple and normally have meanings associated with prototypical property words across languages. They have gradation and are compatible with degree modifiers, as well as intensifying and de-intensifying affixes not available on nouns. Adjectives derived from other word classes are normally non-gradable and employ a distinct derivational morphology.

Adjectives derived from nouns (relational adjectives) typically have the following meanings: “without X, X-less” (privative) (Uyghur *taem* ‘taste’ > *taem-siz* ‘tasteless,’ Kazakh *su* ‘water’ > *su-sız* ‘waterless’), “like X” (comparative) (Chagatay *masa* ‘table’ > *masa-msı* ‘table-like’), “tending towards X, inclined to X” (Uyghur *eve* ‘house’ > *ev-cil* ‘domesticated,’ Kazakh *su* ‘water’ > *su-šıl* ‘capable of swimming in water’), temporal (Sakha *sajin* ‘summer’ > *sajin-ŋi* ‘summer (adjective)’), and locational (Kazakh *iš* ‘inner part’ > *iš-ki* ‘inner’). In some Turkic languages the attributivizing affix can attach to the noun in the locative, that is, the inflectional form, for example Kazakh *új-de-gi*

(house-LOC-ATTR) “located in the house.” Denominal proprietive adjectives are discussed in Section 27.3. Adjectives derived from verbs denote regularity of X (Uyghur *makhtan*- ‘boast’ > *makhtan-chaq* ‘boastful,’ Even *urli*- ‘be jealous’ > *urli-k* ‘jealous’) or the result of X (Turkish *sol*- ‘fade’ > *sol-gun* ‘faded’). Examples of adjectives derived from adverbs are Uyghur *keyin* ‘afterwards’ > *keyin-ki* ‘next’ and Udihe *eineŋi* ‘today’ > *eine-fi* ‘today’s.’

Verbs are a well-defined grammatical class. Inflected verbal words are divided into finite and nonfinite forms. The former may express independent predication, whereas the latter are primarily found in embedded clauses and include participles, converbs, and action nominals/nominalizations. The inflectional categories of finite verbs are subject agreement, tense, and mood. Verbs are formed with a large range of derivational affixes attached to a noun or an adjective (rarely adverbs or minor grammatical classes).

Verbs derived from nouns typically mean “using X as an instrument or means of transportation” (Even *momi* ‘boat’ > *momi-r*- ‘move on a boat’), “possession or acquisition of X” (Uyghur *qoral* ‘arms’ > *qoral-lan*- ‘be armed,’ Even *atikan* ‘wife’ > *atika-lba*- ‘get married (of a man)’), “having something or somebody in the capacity of X” (Even *aman* ‘father’ > *aman-ta*- ‘have as father’), “consuming X” (Even *ulre* ‘meat’ > *ulre-ti*- ‘consume meat’), “providing X” (Even *niki* ‘wild duck’ > *niki-me*- ‘hunt for wild duck’), “smelling of X” (Even *ulre* ‘meat’ > *ulre-m*- ‘smell of meat’), and “acting or being like X” (Mongghul *mongghul-kile*- ‘act like a Mongghul, speak Mongghul’). There are also general verbalizers, whose sole identifiable function is the change of grammatical class, while the resulting meaning depends on the lexical semantics of the base noun and sometimes the context. For example, the very frequent common Altaic suffix *-la* has all these meanings across languages and many more, cf. Kazakh *papa* ‘dad’ > *papa-la*- ‘say “dad”’, *arqa* ‘back’ > *arqa-la*- ‘carry on one’s back,’ *bazar* ‘market’ > *bazar-la*- ‘go to the market,’ *maj* ‘fat’ > *maj-la*- ‘smear with fat’; Udihe *anda* ‘friend’ > *anda-la*- ‘make friends,’ *mä:usa* ‘gun’ > *mä:usa-la*- ‘shoot,’ *zangä* ‘boss, leader’ > *zangä-la*- ‘behave like a boss.’ The same suffix can derive verbs from adjectives, numerals, and adverbs (Uyghur *yakshi* ‘good’ > *yakshi-la*- ‘improve,’ Turkish *temiz* ‘clean’ > *temiz-le*- ‘clean’) and is a component of some morphologically complex affixes (Turkic *temiz-let*- ‘get cleaned’).

Deadjectival verbal suffixes typically indicate acquisition of property (Udihe *baja* ‘rich’ > *baja-na*- ‘become rich,’ Kazakh *boz* ‘pale’ > *boza-ar*- ‘become pale’) or mean something like “to view as X” (Turkish *garip* ‘strange’ > *garip-se* ‘consider strange’). An example of a verb derived from a numeral is Even *gea* ‘second’ > *gea-w*- ‘do something the second time.’ Many Altaic languages employ a distinct derivational morphology in the formation of ideophonic verbs.

27.2.1.2 Category-preserving Suffixation

Evaluative morphology expresses semantic and functional relationships along the axes “small–big” and “good–bad” and does not change the grammatical class of the base, therefore it is discussed in this section. Evaluative suffixes in Altaic are typically external with respect to other derivational suffixes and internal with respect to inflectional morphemes.

Diminutives on nouns are very frequent. Most languages have several, with various degrees of productivity and additional meanings of endearment or contempt. For instance, in Turkish, *-CIK* is the most productive and semantically neutral, *-CAğIz* has an additional meaning of pity or empathy, and *-CAK* additionally expresses endearment but is only present in a few lexical items (Kornfilt 1997: 445). The Even language has the neutral diminutive, diminutive-pejorative, and diminutive-affectionate, and the word can host several diminutives, for example Even *olra* 'fish' > *olra-ča-kan* 'little fish (DIM-DIM)'. Some diminutives only apply to a particular semantic subclass of nouns, for instance, the Kazakh suffixes *-taj*, *-ke*, *-š* and *-žan* are only combinable with human nouns, mostly kin terms or proper names employed in addresses, although Kazakh has a number of other more general diminutives. Augmentatives are mostly typical of Tungusic and Mongolic, and also have additional meanings of approval or disapproval, such as Even *oron* 'domestic reindeer' > *or-kaja* or *oro-nda* 'large domestic reindeer', *oro-mkar* 'good' large reindeer, *bulat* 'fabric' > *bulat-mija* 'old rags'.

Diminutives, often the same as on nouns, also occur on adjectives, for example Turkish *küçük* 'small' > *küçük-cek* 'tiny'. They provide a characterization of the modified noun or indicate the speaker's (affectionate or endearing) attitude towards it, for example Buryat *muu* 'bad' > *muu-xan* 'a little bad (affectionate)', Uyghur *kichik-kinay* 'small (affectionate)'. The moderative suffixes on adjectives denote a small degree of the property, for example Turkish *acı* 'bitter' > *acı-msı* or *acı-mtrak* 'rather bitter', Kazakh *aq* 'white' > *aq-ša* 'whitish', Udihe *sagdi* 'big' > *sagdi-le* 'a little big', Buryat *xara* 'black' > *xara-šag* 'blackish', Dagur *xig* 'big' > *xig-keng* 'rather big', Mongolian *sula* 'weak' > *sula-wtar* 'weakish'. Evaluative verbs are known to be less widespread than evaluative nouns and adjectives. In Altaic they can perhaps include verbs that indicate the degree of intensity of the action. For example, Tungusic languages have intensive verbs, for example Even *nuk-* 'take down' > *nuk-amalči-* 'take down quickly', Udihe *diga-* 'eat' > *diga-ma-* 'eat a lot or eat everything'. In contrast, semelfactive verbs indicate a low degree of intensity, a limited duration or a small amount of the affected object, for example Even *udan-* 'rain' > *uda-sči-* 'drizzle', Udihe *tama-* 'pay' > *tama-nda-* 'pay a little'. Evaluative morphology on the verb does not change its morphosyntactic features or argument structure.

Non-evaluative derivational suffixes that do not change grammatical class exhibit various degrees of productivity, but typically are only restricted to a closed class of lexical items. On nouns they render a large variety of meanings which are difficult to generalize upon. Some are exemplified below: Uyghur *Uyghur* 'Uyghur' > *uyghur-cha* 'Uyghur language', Turkish *ana* 'mother' > *ana-cıl* 'a person dependent on his or her mother', Kazakh *žol* 'road' > *žol-das* 'fellow traveller', Even *nam* 'sea' > *nam-ankan* 'inhabitant of the sea area', *ulre* 'meat' > *ulre-magen* 'the person loving meat', *negdekee* 'bear' > *negdekee-s* 'bear skin', Written Manchu *kumun* 'music' > *kumu-da* 'conductor'. Tungusic languages exhibit a large number of collective suffixes, for example Even *xet* 'willow tree' > *xet-akag* 'willow bushes, willow thicket', *bej* 'person' > *bej-seg* 'family'. For Mongolic one can reconstruct suffixes that derive the names of female beings from tribal names, and age expressions and female animals from color terms (Janhunen 2003: 11).

Deadjectival derivation of adjectives is not very common; some examples are provided by Turkish *som* ‘solid’ > *som-ut* ‘concrete,’ *karşı* ‘opposite’ > *karşı-t* ‘contrary.’ Category-preserving nominal and adjectival suffixes may be borrowed, for instance the Uyghur suffix *-khana* indicating a place connected to the entity denoted by the base noun is borrowed from Persian: *mehman* ‘guest’ > *mehman-khana* ‘hotel.’ The Chagatay *-āna* originates as the Persian adverbial suffix, but in Chagatay it expanded its meaning and can also form adjectives and (by conversion) nouns, for example *turk-āna* ‘Turkish.’ Many Chagatay suffixes of nouns and adjectives go back to Persian verbs, having probably developed from original noun-verb compounds. For example, *-angēz* (from Persian *angīhtan* ‘excite’) has the meaning ‘causing X’ as in *dard-angēz* ‘pitiful, exciting sympathy’ and *niṣat-angēz* ‘causing cheerfulness.’

Verbal category-preserving suffixation involves aspectual, modal and valence-changing processes. Some of these processes are frequent and fully productive, that is, they apply to the whole class of verbs. These instances are not dealt with here because they are fully grammaticalized and may be described as inflection. However, for the most part the relevant processes are restricted to a subset of verbal stems and show semantic idiosyncrasy. Moreover, there may be additional language-specific criteria. For instance, one formal criterion that allows us to distinguish between derivational and inflectional verbal affixes in Tungusic is negation. In periphrastic negative constructions derivational suffixes are hosted by the lexical verb, while inflectional categories are reflected on the negative auxiliary. All derivational suffixes of verbs are easily compatible with each other, and with various inflectional verbal categories.

Aspectual derivation is extremely typical in Altaic, especially Tungusic, producing various imperfective and perfective meanings. For instance, Even exhibits the following imperfectivizing suffixes: durative or continuous (*ma-* ‘kill’ > *ma-d-* ‘be killing’), iterative, frequentative, or multiplicative, which may also denote plurality of subject or object participants (*ma-kat-* ‘kill several objects,’ *ma-gra-* ‘kill several times’), dispersive (*nulge-* ‘roam’ > *nulge-n-* ‘roam from one place to another, roam around’), and habitual (*ma-wat-* ‘kill habitually’). The perfectivizing suffixes are inchoative (*ma-l-* ‘start killing’) and singulative (Udihe *ikteme-* ‘bite’ > *ikteme-le-* ‘bite once’). Other aspect-like derivations typical of Tungusic are repetitive or regressive, for example Udihe *ḡua-* ‘sleep’ > *ḡua-gi-* ‘sleep again,’ *tama-* ‘pay’ > *tama-gi-* ‘pay back,’ and directive, for example Written Manchu *ala-* ‘to say’ > *ala-na-* ‘go to say.’

Modal derivation is less frequent. Some examples are Even *ma-so-* ‘pretend to kill,’ *ma-m-* ‘want to kill,’ *ma-mal-* ‘be ready to kill,’ and *ma-ṣci-* ‘try to kill.’ In Turkish the light verbs *bil-* ‘know,’ *gel-* ‘come,’ *ver-* ‘give,’ and the like can appear as parts of composite suffixes which have various aspectual and modal meanings, for example *yi-* ‘eat’ > *yi-yebil-* ‘be able to eat,’ *gir-* ‘enter’ > *gir-iver-* ‘enter suddenly.’

Valence-changing verbal derivation alters the argument structure of the base verb and typically includes transitivity and detransitivizing processes, such as the causative, decausative, passive, applicative, or the like. Examples of transitivity are Even *xut-* ‘leave’ > *xur-u-* ‘take away,’ Written Manchu *ala-* ‘say’ > *ala-ḡgi-* ‘send to say,’ and

detransitivizers can be exemplified by the Udihe *nodo-* ‘lose’ > *nodo-kpi-* ‘be lost’ and the reciprocalization in Written Manchu *ša-* ‘to see’ > *ša-ndu-* ‘see each other.’ Morphological reciprocals often have the sociative meaning in Tungusic, as in *te-* ‘sit down’ > *te-ndu-* ‘sit down together.’ In Turkic reflexive, passive, and reciprocal suffixes only combine with a limited number of verbs, and the result is often lexicalized, for instance, the Turkish verb *gör-ün-* ‘appear’ is derived from *gör-* ‘see’ by means of the reflexive suffix *-ün-* but it does not have the reflexive meaning ‘see oneself’.

27.2.1.3 Prefixation

As mentioned above, prefixation is not a native morphological process in the Altaic languages. Prefixes are very rare and are only found in languages which exist or existed in the past in a situation of extensive contacts with non-Altaic languages. One such language is Chagatay which borrowed a number of Persian prefixes. A few nouns, particles, and prepositions took on the role of prefixes in classical Persian and were borrowed by Chagatay. They mostly occur in Persian and Arabic nouns, but some spread over to Turkish words. For instance, the Persian prefix *bē-* ‘without’ derives adjectives such as *mihir* ‘sun’ > *bē-mihir* ‘sunless,’ *cān* ‘life’ > *bē-cān* ‘lifeless.’ The prefix *ham-* indicates coexistence or joint agents as in *dard* ‘sorrow’ > *ham-dard* ‘partner in sorrow,’ *rāh* ‘traveller’ > *ham-rāh* ‘fellow traveller.’

27.2.2 Conversion

By conversion or zero-derivation I mean a type of word formation by which a word is shifted into another grammatical class without explicit morphological marking. In many instances the decision on whether we are indeed dealing with conversion is controversial and it is not always clear whether we can talk of a real derivational process. Formally identical nominal and verbal stems are widely spread in Altaic. Janhunen (2003: 10) claims that some stems can actually be reconstructed for Proto-Mongolic as having had both the nominal and verbal function, and the same is true of Turkic and Tungusic. Even if we assume that some kind of zero-derivational process took place, its direction is not easy to establish. One typical assumption is that in a pair “concrete noun–verb” the base word is a noun, while in a pair “abstract noun–verb” the direction is the opposite. In both cases, however, the decision is provisional. Some pairs that can be thought of as involving conversion are Turkish *boya* ‘paint’ ~ *boya-* ‘paint (verb),’ Even *udan* ‘rain’ ~ *udan-* ‘rain (verb),’ and Udihe *bude-* ‘die’ ~ *bude* ‘death.’ Other possible conversional processes are adjective-to-verb (Turkish *kuru* ‘dry’ > *kuru-* ‘become dry,’ Even *berge* ‘fat’ > *berge-* ‘become fat,’ Written Manchu *sagda* ‘old’ > *sagda-* ‘become old’), participle-to-noun (Kazakh *tüis-qan* ‘born at the same time as somebody’ > *tüis-qan* ‘relative’) and adjective-to-noun (Udihe *koto* ‘bold’ > *koto* ‘bold person’), but in fact every major category can be converted into another.

27.2.3 Compounding

Endocentric noun-noun compounds are very frequent. They consist of a nominal head, that is, the categorial part that conveys the basic meaning of the whole compound, and a nominal modifier, which restricts this meaning. Simple noun juxtaposition is widespread, especially in Turkic and Mongolic. An even more productive type involves a compound marker. In Turkic and Tungusic it is homophonous with the third person possessive marker located on the second noun, for example Turkish *okul kitab-ı* ‘school book.’ In Mongolic and Manchu functionally equivalent compounds typically contain the genitive rather than the possessive affix as a compound marker, for example Mongolian *nush-i alcuur* (nasal.mucus-GEN cloth) ‘handkerchief.’ Coordinating compounds or paired words are also widespread.

The noun-verb compounds convert the noun into a verbal structure. The semantics is largely determined by the noun, while inflectional categories are carried by the verb. Such compounds usually contain light verbs and the non-referential object-like noun which stands in the basic form or takes the accusative, for example Turkish *banyo yap-* (bath do) ‘take a bath,’ Udihe *sie-we ba-* (wound-ACC get) ‘get wounded.’ Verb-verb combinations are especially typical of Turkic and Mongolian. In complex predicates the lexical verb determines the primary semantics and argument structure and stands in a nonfinite form. The secondary verb is a semantically light auxiliary which carries all inflection and contributes temporal, aspectual, or modal meaning. Such complex predicates typically denote a single event but the degree of their lexicalization must be empirically decided with respect to each particular language and construction.

27.2.4 Reduplication

Reduplication-like processes are not frequent but do occur. In Turkic and Mongolic there are several common patterns of total or partial reduplication. Intensive/emphatic forms of adjectives, especially color terms, are derived by left-reduplication of the first syllable with the consonant *m*, *p*, *b*, *r*, *n*, or *s* (depending on the language and the phonological context) being added after a vowel-final syllable or replacing the syllable-final consonant, for example Uyghur *qizil* ‘red’ > *qip-qizil* ‘bright red,’ *koek* ‘blue’ > *koep-koek* ‘sky blue,’ Kazakh *boz* ‘pale’ > *bop-boz* ‘very pale,’ *sari* ‘yellow’ > *sap-sari* ‘very yellow,’ *in-ipi* ‘very big,’ Uzbek *butun* ‘full’ > *but-butun* ‘completely full,’ Turkish *temiz* ‘clean’ > *ter-temiz* ‘very clean,’ *boş* ‘empty’ > *bom-boş* ‘utterly empty.’ This is described as prefixation in Kornfilt (1997). In Mongolic this process can be exemplified by Buryat *sagaan* ‘white’ > *sab-sagaan* ‘very white,’ Kalmyk *šar* ‘yellow’ > *šab-šar* ‘very yellow,’ Dagur *xaluun* ‘hot’ > *xab-xaluun* ‘very hot.’

Another common Turkic and Mongolian pattern is the formation of the generic plural ‘all things of the kind; and other things like that.’ This usually involves near-homophonic reduplication where the first consonant of the reduplicated noun is replaced by another consonant: Uyghur *mewae* ‘fruit’ > *mewae-chiwe* ‘all kinds of fruit,’

goesh ‘meat’ > *goesh-poesh* ‘kinds of meat, meat and other ingredients for a dish,’ Turkish *filim* ‘film’ > *filim-milim* ‘film or suchlike,’ Kazakh *nan* ‘bread’ > *nan-pan* ‘bread and the like,’ *šaj* ‘tea’ > *šaj-paj* ‘tea and the like,’ Dagur *xereg* ‘thing’ > *xereg-mereg* ‘all sorts of things.’

Reduplication may involve simple doubling often with an intensifying or adverbial (distributive) meaning, for example Kazakh *žaqsi žaqsi* ‘very good,’ Turkish *usul usul* ‘slowly and softly,’ *damla damla* ‘in drops,’ Written Manchu *ba ba* (place place) ‘everywhere,’ *mejen mejen* (part part) ‘separately.’ In Turkish the particle *-mi* may be inserted between the two components: *güzel mi güzel* ‘remarkably pretty.’ Tungusic languages have adverb-like ideophones underived from a synchronic point of view, although quite a few of them end in *-m*, which suggests that they may have appeared through the conversion of the infinitive of ideophonic verbs not preserved in the modern languages. A large number of ideophones have reduplicated forms, and may sometimes even be repeated three times. The non-reduplicated counterpart may be absent, but in some instances there is a related nominal or adjectival stem, for example Udihe *zogbo* ‘hard’ vs. *zobom zobom* ‘hardly,’ *zūligi* ‘stripy’ vs. *zūle zūle* ‘in a stripy way.’ Some ideophones are derived by reduplication of the final syllable of the base word, cf. *jengute-te* (*jengute-te*) ‘with rumbling’ vs. *jengute-* ‘rumble.’ Reduplication also plays a role in the onomatopoetic vocabulary in Mongolic, for example in Dagur (Janhunén 2003: 135).

27.3 PROPRIETIVE ADJECTIVES

Dividing derivational processes into two types depending on whether they change grammatical class is not straightforward. The so-called “mixed categories” are problematic for the traditional classification of syntactic categories, as they exhibit properties of more than one part of speech: phrases of two different syntactic categories appear to be projected by the same word. Such cases are well known from the vast literature on nominalizations, but Altaic languages possess other types. This section presents a case study of proprietive adjectives productively derived from nouns.

27.3.1 Fully Adjectival Proprietives

Altaic languages exhibit a large number of adjectives whose meaning can loosely be described as proprietive, that is, “with X” or “having X,” where X is the base noun. Proprietives fall within a large semantic range, but, most generally, they can be characterized as denoting possession, if the latter is understood fairly loosely as a certain (pragmatically and/or lexically defined) association between two entities. But when the proprietive adjective modifies a noun, the semantic relation is reversed compared to the possessive construction: the head noun is the “possessor,” while the dependent adjective

corresponds to the “possessed,” for example *rainy autumn* (= *autumn with rain*), *hungry dog* (= *dog with hunger*), *bearded man* (= *man with beard*).

Proprietary adjectives in Altaic are derived by suffixation, and for the most part the suffixes are rather old and can be reconstructed for the respective proto-families. For instance, the most common Turkic proprietary suffixes *-lu* and *-lig* go back to Proto-Turkic, and for Mongolic the proprietary morpheme is reconstructed as *-ta(y)i* (Janhunen 2003: 11). However, there are also instances of borrowing: the Chagatay proprietary suffix *-dār* as in *āb* ‘lustre, splendor’ > *āb-dār* ‘brilliant’ goes back to the Persian verb *dāštan* ‘to hold.’

Some fully adjectival proprietives are derived by the same suffixes as proprietives with the nominal properties addressed in the next subsection. However, they have fully lexicalized as regular adjectives and do not exhibit any of the characteristics of nouns, as is typical of the “dual” proprietives. Adjectival proprietives form a closed class and usually denote a more or less permanent property, for example Evenki *saŋaar* ‘hole’ > *saŋaa-gde* ‘with holes,’ Udihe *zolo* ‘stone’ > *zolo-n’ouse* ‘stony,’ *guakta* ‘beard’ > *guakta-n’ause* ‘bearded,’ Written Manchu *mambi* ‘knot’ > *mambi-ŋe* ‘with knots,’ Buryat *gerge(n)* ‘wife’ > *gerge-tej* ‘married.’ If the proprietary is derived from an abstract or mass noun, the original proprietary meaning is only preserved metaphorically; the adjective simply denotes an abstract feature: Sakha *küüs* ‘strength’ > *küüs-teex* ‘strong,’ Chagatay *rang* ‘color’ > *rang-lig* ‘colorful,’ Kalmyk *max* ‘meat’ > *max-lyg* ‘fleshy,’ Buryat *torgon* ‘silk’ > *torgo-lig* ‘silky,’ Even *nes* ‘luck’ > *nes-eč* ‘lucky,’ Written Manchu *doro* ‘law’ > *doro-ŋe* ‘legal,’ Nanai *murun* ‘mind’ > *murun-ku* ‘intelligent.’ This can make proprietary adjectives somewhat ambiguous; for instance, the Buryat *bešeg-tej* derived from *bešeg* ‘letter’ has both the literal meaning ‘having a letter, with letter’ and the more abstract meaning ‘literate.’ In some instances the same proprietary affixes indicate a large quantity of the entity denoted by the base noun (Buryat *dobo* ‘hill’ > *dobo-lig* ‘having many hills,’ Even *oron* ‘reindeer’ > *oro-pči* ‘having many reindeer’) or the large size of this entity (Evenki *ije* ‘horn’ > *ije-regde* ‘with long horns,’ Even *amŋa* ‘mouth’ > *amŋa-gda* ‘with big mouth’).

In some languages the proprietary suffixes have more diverse functions. They may be employed in deverbal derivation, for example Written Manchu *tat-* ‘pull’ > *tata-ŋe* ‘stretched, pulled out.’ They may develop secondary meanings such as “similar to X” (e.g. Buryat *morin* ‘horse’ > *mori-rxu* ‘similar to a horse’), “located in X” (Buryat *ajaga* ‘cup’ > *ajaga-taj* ‘having a cup; being in a cup’) or “belonging to a place or institution” (Turkish *Ankara* ‘Ankara’ > *Ankara-lı* ‘person living in Ankara’). Proprietary adjectives often undergo substantivization, for example Kazakh *mal* ‘cattle’ > *mal-dı* ‘having (a lot of) cattle’ > ‘collective farm,’ Kalmyk *tsetseg* ‘flower’ > *tsetseg-lg* ‘with flowers’ > ‘park.’

Adjectival proprietives function as modifiers or predicates, but never as independent arguments or adjuncts. The most important property that allows classifying them as adjectives is their prenominal position within a noun phrase and the absence of any marker of syntactic dependency (such as, for example, the genitive case or possessive

agreement). But in Evenki and Even, which have attributive concord, adjectival proprieties agree with their heads in number and case, cf. the following example from Even:

- (4) taŋn 'a-gda-w čalam-u
 antlers-PROPR-ACC reindeer-ACC
 'the reindeer (ACC) with big antlers'

At least some of the adjectival proprieties may take morphological markers of the comparative degree (Cincius 1947: 110). The base noun of adjectival proprieties does not retain any referential characteristics: it cannot be cross-referenced by anaphoric elements, does not pluralize and does not head a syntactic phrase of its own. As derivation is usually thought of as having no projection into clausal syntax, these facts classify the resulting forms as derivationally formed adjectives.

27.3.2 Proprietary Adjectives with Nominal Properties

In this section I address the proprieties which, on the one hand, behave like regular adjectives, and, on the other hand, have certain properties of nouns, unlike the fully adjectival proprieties described in the previous subsection.

Most importantly, the base noun from which the propriety is derived can head its own syntactic phrase, that is, take its own modifier. This is illustrated for Turkic and Mongolic by the Chagatay *aq saqal-li* (white bear-PROPR) 'with a white beard,' *qaba yüz-lüg* (ugly face-PROPR) 'with an ugly face,' Turkish *iki litre-lik kove* (two litre-PROPR bucket) 'two-litre bucket,' *sarı saç-lı kız* (blond hair-PROPR girl) 'blond-haired girl,' Salar *tö'tur tumzux-li* (four angle-PROPR) 'with four angles,' and Kalmyk *xoyr niilwr-tä ügmüd* (two syllable-PROPR word.PL) 'bi-syllabic words.' However, in Turkic and Mongolic modification is generally restricted to numerals and qualitative adjectives in constructions where the relation between the entities denoted by the head noun and the nominal base of the propriety is that of inalienable possession (typically body parts as in 'blond-haired girl'), rather similar to *-ed* adjectives in English. In Tungusic, on the other hand, there are no such restrictions: the propriety base can be modified by an adjective independently of its semantics, as shown by the Udihe *aja mä:usa-xi* (good gun-PROPR) 'with a good gun,' *xulaligi xei-xi* (red handle-PROPR) 'with a red handle' and Evenki *n' amapču aaw-laan* (warm hat-PROPR) 'with a warm hat.' Moreover, the nominal base can head a quantifier (Nanai *xaidaxem xadon-ko* (all.sort equipment-PROPR) 'with all sorts of equipment'), a reciprocal adjective (Udihe *mene-mene eze-xi* (REC-REC master-PROPR) '(each) with his own master,' a nominal modifier (Udihe *bui je-xi* (deer antlers-PROPR) 'with deer antlers'), an oblique phrase (Nanai *ŋala-gdo-i gida-ko* (hand-DAT-REFL spear-PROPR) 'with a spear in his hands'), or an apposition-like noun (Evenki *Tokko gerbi-či bira* (Tokko name-PROPR river) 'the river named Tokko'). The only types of adnominal dependents that seem to be impossible for the

base noun are possessors and determiners, so we may say that proprietive derivation involves attributivization of NPs but not DPs.

Tungusic proprietives have other nominal properties. The base noun may show number opposition, and in Northern Tungusic, which has attributive concord, it triggers number concord on its modifier, as shown in (5b).

- (5) a. *danša-sal-ku* *tetredi-sal-ku* *nuchikend'uen* Nanai
 book-PL-PROPR notebook-PL-PROPR children
 'children with books and notebooks'
- b. *xulaña-l* *hamu-r-alkar* Even
 red-PL flag-PL-PROPR
 'with red flags'

The underlying noun is available for cross-reference by anaphoric elements or the referential null. In examples (6) in Udihe the proprietive adjectives function as predicates; the base noun "sister" is referred to by the anaphoric element *uti* in the following clause, while the base noun "husband" functions as a pronominal null subject.

- (6) a. *Uti* *zu* *ni:* *xun'azi-xi* *bi-si-ti*
 that two man elder.sister-PROPR be-PAST-3PL
 Uti 'ana-du *bi-si-ni*
 that boat-DAT be-PAST-3SG
 'These two men had an elder sister. She was in the boat.'
- b. *Mafa-xi* *bi-de* *bi-si* Ø *emuge-we*
 husband-PROPR me-FOC be-PAST cradle-ACC
 o:-muse
 make-COND
 'If I had a husband, he would have made a cradle.'

So the base nouns have many properties of independent referential nouns. This has a semantic motivation: unlike the fully adjectival proprietives discussed in the previous subsection, the proprietives addressed in this section are typically derived from concrete countable nouns and denote temporary possession rather than a permanent characteristic. Thus, the Nanai proprietive construction *mogo-ku koaŋsakan* (mushroom-PROPR basket) can only have the meaning 'basket with mushrooms, i.e. basket that has mushrooms inside it,' but cannot mean 'basket used for mushrooms.' To render the latter meaning a compound must be used.

According to the criteria mentioned in Section 27.2.1.1, the proprietive forms addressed in this subsection should be classified as adjectives. They can only function as adnominal modifiers, secondary predicates, or main predicates, but do not occur as independent arguments or adjuncts. In the modifying function they are always pronominal and show attributive concord with the head in Northern Tungusic, as do other adjectives, see the following Evenki examples:

- (7) a. *ŋinaki-či-l-va* *beje-l-ve*
 dog-PROPR-PL-ACC man-PL-ACC
 ‘men (ACC) with dogs’
- b. *oro-či-l-du* *asa-l-du*
 reindeer-PROPR-PL-DAT woman-PL-DAT
 ‘to the women with reindeer’

In this they differ from purely nominal categories with proprietive-like meanings, usually described as case forms of nouns (comitative or instrumental). These can also function NP-internally, but their position in the NP differs: in Tungusic they typically stand in postposition to the head, for example Even *asi kuŋa-n’un* (woman child-COM) ‘a woman with a child,’ Evenki *bejumimni pektyrevu-tei* (hunter gun-COM) ‘a hunter with a gun.’ Like adjectives and unlike nouns, the proprietives in question can function as secondary predicates with verbs of motion and verbs of creation. It should also be added that the relevant proprietives can be modified by (at least some) adverbs, as is typical of adjectives but not of nouns:

- (8) a. *emereč* *tobar-alkan* *molimŋa* Even
 sharply axe-PROPR wood.cutter
 ‘wood-cutter with a sharp axe’
- b. *padi-padi* *omo-ko* Nanai
 separately-separately hole-PROPR
 ‘with separate holes.’

More adjectival properties are addressed in Nikolaeva (2008), but in sum, the proprietives in question pass all the adjectivehood tests relevant for the Tungusic languages.

So if nouns and adjectives are to be distinguished in Tungusic, proprietives should definitely be classified as adjectives. This implies that they are derived from nouns by an interclass derivational process. However, with the assumption that derivation is opaque to syntax, we cannot account for the active syntactic properties of the base noun. On the other hand, the analysis of the relevant proprietives as inflectional forms of nouns leaves their adjectival properties largely unexplained.

A further problem is the morpheme order. Derivational morphemes are usually assumed to be located closer to the root than inflectional morphemes. However, as shown in (5), the inflectional plural precedes the proprietive marker. This is also the case with the Tungusic so-called inalienable possessive affix *-ŋi/-ŋo*. This affix must be combined with a regular possessive affix on the noun that heads a possessive phrase. Its function is to indicate that the relationship between the possessor and the possessee is not a prototypical possessive relationship, but some kind of (sometimes rather loose) association established depending on the context. With this function the “alienable possessive” suffix can be hosted by any noun and thus could be regarded as an inflectional marker. This suffix is compatible with the proprietive morpheme. Crucially, like the plural suffix, it regularly precedes it, for example Udihe *belie-ŋi-xi* (fairy-AL-PROPR) ‘with a fairy’

and Nanai *gasa-ŋo-ku* (duck-AL-PROPR) ‘with a duck.’ If the morphotactic requirement mentioned above is to be taken seriously, this too presents a problem for interpreting propriative formation as a derivational process. On the other hand, the evaluative suffixes characterizing the base noun can be external to the propriative morpheme, as in the Evenki *saŋari-či-kaar* (hole-PROPR-DIM) ‘with little holes’ and *dere-či-keekuun* (face-PROPR-AUGM) ‘with a big face.’ This presents a problem for analyzing propriatives as some kind of phrasal affixes.

We are left with the conclusion that the categoriality for such propriatives cannot be unambiguously established. They demonstrate mixed behavior in that they have both adjectival and nouny properties. The adjectival properties are categorial, that is, those that characterize an adjective as an adnominal modifier (the external syntactic behavior). The nouny properties are mostly internal, that is, they have to do with the compatibility of the base noun with other morphological and syntactic elements within their own phrase. These properties characterize a noun as a discourse referent.

Finally, it should be noted that there is a historical connection between propriative adjectives and the comitative forms of nouns. Comitatives are inflectional case forms which in Altaic can function as clausal arguments and adjuncts and are sometimes used for coordination. The Mongolic propriative suffix *-*tA(y)i* mentioned above derives a propriative adjective, but it is also found on nouns that function at the clause level and is described as a comitative for some languages. As the examples in (10) show, unlike “dual” propriatives, comitative nouns are fully nominal and compatible with possessive morphology.

- (9) a. *eshkee-tii-miny* *eus* *lashiiciseng* Dagur
 uncle-COM-1SG grass went.to.cut
 ‘He went to cut grass with my uncle.’
 b. *mori daaha-dii-naa* *fugue xija* Mongghul
 horse foal-COM-REFL dies
 ‘The horse dies together with its foal.’

The direction of historical development (from adjectival suffix to the comitative case, or the other way round) is not entirely clear, but the historical connection between the two suggests that various intermediate stages and forms with mixed properties are likely to arise.

CHAPTER 28

YENISEIAN

EDWARD J. VAJDA

THE Siberian microfamily Yeniseian contains the extinct Yugh, Arin, Pumpokol, Assan, and Kott (Werner 2005), as well as the severely endangered Ket language, spoken fluently by no more than a few dozen elders in isolated villages near the Yenisei River. The family's two best attested primary branches, Ket-Yugh and Kott-Assan, are fairly closely related and probably separated no earlier than 2500 years ago. The scantily attested Arin and Pumpokol may also have been primary branches, though a few features of Arin verb structure appear to be innovations shared with Ket-Yugh. Yeniseian grammatical morphology consists of possessive prefixes and case suffixes on nouns, postpositions joined to nouns by possessive morphology, and a strongly prefixing finite verb that differs strikingly from all other word types, except insofar as infinitives (action nominals) are derived from it. Morphologically complex noun stems are built mainly through root compounding, but also involve a few extremely productive derivational suffixes. Finite verb morphology is templatic and largely prefixing, with lexical and grammatical morphemes intertwined in linearly complex patterns. Ket, with three surviving dialectal forms,¹ is the family's best-described member (Werner 1997c, Vajda 2004, Georg 2007). Yugh died out in the 1970s and was also well documented (Werner 1997b). Kott was last described in the 1840s on the eve of its extinction but was much better recorded than the other extinct southern languages (Castrén 1858, Werner 1997a, 2005). Phonemic tones, possessive prefixes, and a templatic prefixing verb are Yeniseian typological features that differ starkly from other North Asian families (Vajda 2007). Morphological and lexical evidence supports a distant genealogical connection with the Na-Dene family of north-western North America (Vajda 2011).

¹ Examples are Southern Ket except where otherwise noted. The Northern, Central, and Southern Ket dialects differ in some aspects of phonology and vocabulary and are mutually intelligible.

28.1 THE TYPOLOGY OF YENISEIAN WORD FORMATION

The Yeniseian lexicon can be divided into three basic morphosyntactic classes: nouns, modifiers, and finite verbs. Nouns (defined as forms capable of taking possessive prefixes and case suffixes) are morphologically of three types. Many basic nouns are non-derived monosyllables such as *ke't* 'person', *ā·m* 'mother', *tī·k* 'snow', *kūn* 'wolverine', *su:l* 'sled', *sē·s* 'river', *qo'n* 'conifer branch', *dī·k* 'resin', *qu'j* 'birchbark'.² Other noun stems contain what could be described as prefixes or suffixes, though nearly all such nominal affixes appear to have originated from compounded noun roots. The suffix *-kit* (< *ka'd* 'offspring') derives nouns denoting immature trees, fish, or animals: *be's* 'rabbit' > *beskit* 'baby rabbit', *qu'r* 'pike (fish)' > *qurgit* 'immature pike', *u:l* 'aspen' > *u'lgit* 'aspen sapling'. The singulative suffix *-dis* (< *dē·s* 'eye') derives count nouns denoting a small, individuated portion of a larger pourable mass: *imdis* 'single pine nut', *qondis* 'glass bead', *uldis* 'water droplet' (< *i·m* 'pine nuts', *qō·n* 'beads', *ū·l* 'rain, water, liquid'). Less easy to detect as an underlying root is the suffix **-p^hul* 'growth, mass, accumulation,' found in Ket and Yugh as the second element in a number of count nouns: Ket *aspul*, Yugh *asful* 'cloud' (< **ē·s* 'sky' + **p^hul* 'mass'), Ket *nanul* ~ *nanl*, Yugh *n'enjful* 'bun, roll' (< **n'a'n* 'bread' + **p^hul* 'mass'). This element is not found as a free root in Ket or Yugh, yet has a cognate in the Kott verb root *-fel* 'grow'. The feminine derivational suffix *-a* found in Kott and Ket-Yugh, which would appear to be a purely derivational element, probably originated from parallel borrowing of the Russian feminine-gender inflection *-a*. Rather than shared between Ket–Kott cognates, it appears in different lexemes, such as Kott *dā:le-a* 'girl' (< *dā:le* 'child') or Ket *Sala*, a female personal name derived from the male name *Sa'l* meaning 'sharp edge'. Many Arin loanwords end in the syllable *-ok* (Werner 2005: 149), which resembles the Russian diminutive suffix *-ok* that may have spread analogically to mark all loan vocabulary: Arin *gorox-ok* 'peas', *kuris-ok* 'hen' (< Russian *gorox* ~ *gorošek* 'peas', *kuritsa* 'chicken'), also Arin *kajak-ok* 'butter', *altin-ok* 'gold', *kumiš-ok* 'silver' (cf. Old Turkic *kan'ak* 'clotted cream', *altuñ* 'gold', *kūmuš* 'silver').

The fact that most Ket noun affixes retain an etymological connection to noun roots suggests that even nouns built through productive derivational processes originated as root compounds structured semantically as modifier + head. One particularly clear

² Southern Ket monosyllables have four phonemic tones: *ṽ·*—high-even tone on a half-long vowel; *v'*—abrupt rising tone on a short vowel ending in laryngealization or optionally with a glottal stop; *v*—rising-falling tone on a geminate vowel; *ṽ* falling tone on a short vowel (Yugh has a long pharyngealized vowel in falling tone words: *ṽ^h*). There are seven vowel phonemes: *i*, *ε*, *a*, *ɔ*, *ʌ*, *u*, *u*. The vowel *u*, often transcribed by Ketologists as *i*, is a high back unrounded vowel (the phonetic equivalent of *i* in Turkological transcriptions). The mid-vowels *ε*, *ɔ*, *ʌ* have raised allophones under high-even tone: *ē*, *ō*, *ṽ*. The vowel phoneme *ɤ* ~ *ʌ* has often been transcribed as *ə* though it is mid-back rather than mid-central. Ket multisyllabic words usually have a rising-tone peak on the first syllable, less often on the second—both tending to become non-tonal word stress in connected speech (see Vajda 2004, Georg 2007).

example is *-oks* (< *ō'ks* 'tree, stick'), a suffix found in many nouns expressing long wooden objects: *bejoks* '(ship's) mast' (*bē'j* 'wind'), *tunteroks* 'stick used to hang a kettle over the fire' (*tu'n* 'kettle', *ted* 'across'). Another is *-am* (< *qī'm* 'woman') in *ennam* 'daughter-in-law', *qannam* 'tsarina', *jukkam* 'Yugh woman' (cf. *ēn* 'son-in-law', *qā'n* 'khan, tsar', *jūy* 'Yugh person'). Yet another is the generic body-part suffix *-at* (< *ad* 'bone'): *kil-at* 'hip, thigh', *lan-at* 'hand', *ul-at* 'rib', *qov-at* 'back' (apparently found also in Kott *ul-aj* 'rib', *hap-ar* 'back'). Overall, there are about a dozen noun-deriving suffixes of this sort, which Werner (1997c: 50–4) aptly calls 'half-affixes' to underscore their transitional nature between compounded root and derivational affix. A few serve as the first, or modifying element of the compound: *hanrip* 'female dog', *hanrit* 'female spruce grouse' (< *hàn* 'female animal' + *tīp* 'dog', *dī't* 'spruce grouse'); *uktip* 'male dog', *ugdit* 'male spruce grouse' (< *u'k* 'male animal').

Among noun-deriving elements, only the extremely productive nominalizer Ket *-s*, Yugh *-su*, Kott *-še* does not seem to derive from an earlier noun root (its origin is unclear). It creates a large, in fact potentially open-ended, number of nominalizations (discussed in Section 28.2). The remaining Ket nouns are root compounds based, with varying degrees of transparency, on the same model of modifier + head described above in connection with the Ket-Yugh half affixes: *higdul* 'boy' (*hī'y* 'male', *dū'l* 'child'), *untip* 'puppy' (*hana* 'little', *tīp* 'dog'), *kanqo* 'Venus' (*ka'n* 'bright', *qōχ* 'star'), *ē'y su:l* 'automobile' (*ē* 'iron', *su:l* 'sled'). Certain possessive constructions have become lexicalized as nouns: *destuj* 'eyeball' (lit. 'eye's belly' < *dē's* 'eye' + *di* 'inan.-class poss.' + *hū'j* 'belly'); *destqar* 'eyebrow' (lit. 'eye's fur' < *dē's* 'eye' + *di* 'inan.-class poss.' + *qā'r* 'fur'); *mōrat bō'k* 'aurora borealis, northern lights' (lit. 'sea's fire' < *mōra* 'sea' + *di* 'inan.-class poss.' + *bō'k* 'fire'), *ekkanma qō't* ~ *aqqot* 'rainbow' (lit. 'path of thunders' < *eknan* 'thunders' + *na* 'animate-class plural poss.' + *qō't* 'path'). Modifier + head compounds, including half affixes and lexicalized possessive constructions, when considered together with nouns derived by nominalizing *-s*, account for nearly all morphologically complex noun stems.

Modifiers include adjectives, adverbs, numerals, as well as infinitive forms (to be discussed in Section 28.4). As a morphological category, modifiers can be defined as any content word that is not a finite verb and which cannot take inflectional suffixes without first being nominalized by the universal nominalizer *-s*. Modifiers used predicatively normally take subject agreement suffixes, whereas nouns do not: *ā't ō'p* 'I am a father' (I + father), but *ā't aqta-di* 'I am good' (I + good-1SG.PRED.AGR.), *ā't kisēn-di* 'I am here' (I + here-1SG.PRED.AGR.). Adjectives may be simple roots (*ki* 'new', *aqta* 'good', *sē'l* 'bad') or contain a fossilized adjective suffix *-aŋ*, as in *hilaŋ* 'sweet' < *hi'l* 'birch sap' + *aŋ* (adj. suff.). This suffix may have been more widespread in the past, as a couple of dozen Ket adjectives seem to have reanalyzed *-aŋ* as the homophonous nominal plural marker, so that it has come to be used in conjunction with plural nouns, while elsewhere disappearing or reducing to falling tone (less often to another tone): *qā* ~ *qāŋ* 'big'; *bō'l* ~ *bōlaŋ* 'fat', *ugde* ~ *ugdeŋ* 'long'. This reanalysis explains an otherwise idiosyncratic feature in a strongly head-marking language like Ket (Vajda 2013).³ Denominal adjectives conveying the underlying noun's presence or

³ Vajda (2004: 80), following earlier Ketologists, described this suffix as a plural agreement marker, additionally suggesting it might mark "tangible physical qualities." Reanalysis of an original adjective suffix that elsewhere tended to erode is a more likely explanation.

quality are productively formed by suffix Ket *-tu*, Yugh *-čoun* (the latter probably retaining the original adjective suffix **-əŋ*): Ket *qim-tu*, Yugh *χem-čoun* ‘with a wife, married (said of a man)’; Ket *sul-tu*, Yugh *sur-čoun* ‘bloody’; Ket *tas-tu*, Yugh *čas-čoun* ‘stony, full of rocks’.

Some adverbs also consist of what appears to be a simple root (*i’n* ‘for a long time,’ *ən* ‘many’). Temporal nouns can function as adverbs of duration without morphological modification (*sī* ‘night,’ ‘at night’; *qogdi* ‘fall, autumn,’ ‘in the fall’). Adverbial suffixes include **-p^ha*, added to numerals to denote number of repetitions: Ket *una*, Yugh *unfa* ‘twice’; Ket *donə*, Yugh *donfa* ‘thrice’. There is also the distributive suffix Ket *-sa*, Yugh *-so* ‘each, every,’ probably derived from the Proto-Yeniseian root **suk* ‘back, return’: Ket *qonoks-sa* ‘every morning,’ *bis-sa* ‘every evening,’ *is-sa* ‘every day,’ Yugh *i.^h-so* ‘daily’. Many nouns with relational suffixes such as the caritive suffix *-an*, which denotes absence, serve either as adjectival modifiers (*qim-an ke’t* ‘bachelor,’ lit., ‘wife-less man’) or verbal complements (*qim-an duyaraq* ‘he lives without a wife,’ lit., ‘wife-without he lives’). In this way, the numerous relational morphemes postposed to nouns (case endings and postpositions) fulfill the derivational function of creating modifiers by negating the noun’s syntactic head status (Vajda 2008). Many adverbs derive in this way. The prosecutive suffix *-bes*, meaning ‘passing by, along, through,’ is added to plural nouns of nationality such as *əstuyan* ‘Ket people’ (< *əstuk* ‘Ket (Ostyak) person’) to denote communication in a language: *əstuyan-bes* ‘in Ket,’ *kusn-bes* ‘in Russian,’ *layan-bes* ‘in Selkup,’ *hamban-bes* ‘in Ewenki,’ etc.

Finite verbs differ morphologically from nouns and modifiers in being formed on the basis of a complex position class template. Templatic interrelationships between different verb stem patterns in Modern Ket are discussed in Section 28.3. Every conjugated verb form is polymorphemic and some fill as many as eight distinct affixal slots. The verb morphology of Yugh and Kott is similar, except that Kott innovated a new subject agreement suffix position at the very end of the verb. Because subject markers appear in most conjugated verbs, Kott superficially appears more strongly suffixing that it actually is. In fact, Kott inherited seven prefixal positions from Proto-Yeniseian—positions likewise inherited by Ket-Yugh as slots P7–P1 in the model shown in (3) in Section 28.3. So few Arin and Pumpokol verb forms were recorded that the relative prevalence of prefixing vs. suffixing in these languages is difficult to assess. At least some Arin verb forms (*ba-x-a-tum* ‘I go’) show prefixal strings that are clearly homologous with those of Ket (*bɔ-y-a-tɨn* ‘I go’) and Yugh (*bɔ-a-de* ‘I go’)—all presumably from Proto-Yeniseian or Proto-Ket-Yugh-Arin **ba-g-a-den^j* 1SG.SBJ-away-PRES-go.

Affixation or cliticization⁴ in Yeniseian is predominantly inflectional. Both types of syntactic heads—nouns and finite verbs—take many inflections and other bound relational morphemes. Nominal forms permit only a single possessive prefix (actually a clitic), but may take a concatenation of plural, case, and postpositional suffixal elements. The complex denominal modifier in (1) contains the postposition *bal* ‘between, among,’ which arose through grammaticalization of an earlier noun root meaning ‘gap, space’ attached to the preceding noun by the use of possessive morphology. Postpositional constructions

⁴ Vajda (2008) argues that most nominal inflections are actually clitics rather than affixes.

are completed by a case marker such as the possessive + ablative combination *-di-ŋal* used in (1a) to signify ‘from,’ or locative *-ka* used in (1b) to signify stationary location:⁵

- (1) a. *da-huub-aŋ-na-bal-di-ŋal*
 FEM.SG.POSS-SON-PL-ANIM.POSS.PL-between-INAN.POSS-ABL
 ‘from between her sons’ (literally, ‘her-sons-’-gap-’s-from’)
- b. *b-huub-aŋ-na-bal-ga*
 1SG.POSS-SON-PL-ANIM.POSS.PL-between-LOC
 ‘located between my sons’ (literally, ‘her-sons-’-gap-in’)

Finite verb stems, in contrast to the other form classes, are constructed based on a template arrangement in which the derivational and inflectional material is syntagmatically interdigitated, creating a complex morphological string that contrasts sharply with nominal morphology (see Section 28.3). Incorporation (technically a type of compounding) is restricted to a few verb stems, while noun stem formation is frequently based on modifier + head compounds. The example in (2) shows a verb with an incorporated noun object:

- (2) *k-ul-a-dɔp-in*
 2SBJ-water-PRES-drink-ANIM.PL.SBJ
 ‘you (PL) drink water’

Only a few basic verb roots permit object incorporation; the majority of transitive stems do not allow it. Example (2) also illustrates how verb forms intersperse agreement and tense–mood inflections between positionally disjunct lexical morphemes. The model in (3) in the next section will provide more information about finite verb structure.

Yeniseian derivational affixes are less numerous than inflectional affixes and clitics, even when the so-called half affixes are included. Very few derivational suffixes, perhaps only nominalizing *-s* and adjectival *-tu*, are fully productive. There is no reduplication in either inflection or derivation. Ket is a strongly synthetic language in terms of its morphosyntax but its stem-building processes are based primarily on root compounding rather than affixation.

28.2 NOMINALIZATION

Virtually any modifier or grammatically inflected word form in Modern Ket can be nominalized by adding the suffix *-s*: *kʌn* ‘bright, light colored’ > *kʌns* ‘one that is bright,

⁵ Three case endings (ablative, dative, adessive) require a possessive connector, while the others (locative, prosecutive, instrumental, caritive) do not. The reason for this pattern is inexplicable from the semantics and phonology of the modern language. Denominal postpositions (like *bal* ‘between’ < ‘gap’) also require a possessive connector. See Vajda (2008).

something shiny, someone light-skinned'; *sultu* 'bloody' > *sultus* 'something (or someone) that is bloody'; *tasthitiya* 'under a rock' > *tasthitiyas* 'one that is under a rock'; *qiman* 'wifeless' > *qimans* 'one that is wifeless, bachelor'; *kisɛŋ* 'here' > *kisɛŋs* 'one that is here'; *ɛŋŋunɬɐr* 'to build houses, house-building' > *ɛŋŋunɬɐrs* 'one that builds houses, carpenter', etc. Some nominalizations in *-s* have acquired idiosyncratic meanings, such as *qà* 'big' > *qàs* 'boss' (alongside the original meaning 'one that is big') and possibly *kuuns* 'Russian' < *ka'n* 'bright, light-colored', referring to pale skin. Nominalizations in *-s* can take possessive prefixes or case suffixes just like any other noun. Also like other nouns, they can appear as predicate nominal without the subject agreement suffixes typical of predicate adjectives and adverbs: *ke't qa-ru* 'the person is big' (with masculine-class singular predicative agreement suffix *-du*), but *ke't qa-s* 'the person is a big one' (with nominalizing *-s*). In rare cases nominalizing *-s* is added to a noun used predicatively to derive a more abstract meaning: *qī'm* 'woman' > *qim-s* 'is a female' (Werner 1998: 61). The resulting nominalizations in *-s*, however, cannot serve as attributive modifiers, though the bare noun itself may serve as modifier in root compounds: *qimduul* 'girl' < *qī'm* 'woman' + *dū'l* 'child'. Their inability to serve as attributive modifier to another noun is the one trait that functionally distinguishes nominalizations in *-s* from other nouns.

The universal nominalizer can also be added to any finite verb forms to create a noun or headless relative clause: *dɔldaɔ* 'he lived' > *dɔldaqs* 'one who lived', *dbanbun* 'they repeatedly give me something' > *dbanbuns* 'those who repeatedly give me something', *dbanbilun* 'they used to give me something' > *dbanbiluns* 'those who used to give me something', *daqalajuksej* 'it has turned yellow' > *daqalajuksejs* 'one that has turned yellow', etc. Nominalizing *-s* is the only affix capable of nominalizing a conjugated verb form. Basically any past or present indicative verb can be nominalized in this way (though imperatives cannot). The *-s* suffix cannot appear if the verb form is used as an attributive modifier to another noun: *dɔldaqs* 'one who lived, he who lived' (< *dɔldaɔ* 'he lived'), but *dɔldaɔ ke't* 'the man who lived'.

28.3 TEMPLATIC DERIVATION

Yeniseian finite verb forms are based on a rigid position class model in which the semantic head (the verb root) originally occupied a position near the verb's rightmost edge, so that most of the additional morphological material appeared as prefixes. The model in (3) is arranged so that the slot shown at the top (the only suffixal position in Modern Ket-Yugh, and also inherited from Proto-Yeniseian into Kott) represents the rightmost morpheme in the verb form.

(3) Modern Ket finite verb structure

P-1	Animate-plural subject suffix ⁶
Po	Base (verb root functioning as the semantic head of simple verbs, but reduced to a semantically bleached form serving as aspect or voice marker in verbs with a salient P7 incorporate, which now account for all productive ways of forming Ket verbs)
P1	Original speech-act-participant subject agreement position (now filled only in certain stative or anticausative verbs, and used instead for object marking in many other stems)
P2	Lexical aspect marker in most past-tense indicative or imperative forms (generally <i>l</i> -imperfective, <i>n</i> -perfective)
P3/4	Conjugation prefix (<i>si-</i> , <i>i-</i> , <i>a-</i> , <i>ɔ-</i> , rarely <i>qɔ-</i>) used to express tense and mood in combination with the P2 aspect markers <i>l</i> and <i>n</i> . 3rd person agreement marking (originally subject markers, now object markers in some stems) originally occupied the preceding position P4. However, inanimate-class <i>b-</i> has metathesized rightward ahead of the conjugation marker in Ket-Yugh, complicating the original linear order of P3 conjugation marker preceded by P4 agreement marker (an arrangement retained unaltered in Kott).
P5	The so-called ‘determiners’—obligatory stem elements consisting of a single consonant; lexical morphemes that express shape or trajectory, but more often are semantically opaque. Determiners sometimes concatenate in a single verb form.
P6	The original object agreement position, now used in many stems for marking subject agreement instead.
P7	The incorporate position, which may contain an incorporated noun, adjective, directional adverb, or infinitive. In verbs with P7 infinitives, the Po base is reduced to an auxiliary function (inception, voice marking, iterativity, etc.).
P8	The new subject person agreement clitic in Modern Ket (and Yugh). Filled in most action verbs or resultative verbs to express agreement with animate-class subjects (also, more rarely, with highly active inanimate-class subjects). This slot never developed in Kott, which instead innovated a new suffixal subject agreement position (P-2).

A Ket verb stem can be represented as a formula composed of the lexical morphemes from P7, P5, Po, plus the lexical choice of which tense–mood–aspect affixes from P4 and P2 the stem will use, and finally the lexically idiosyncratic choice of which slots will be used to express subject and object agreement. Ket has six productive tense–mood classes, five productive intransitive configurations, and three productive

⁶ Found in all verbs that use P8 for subject person agreement, except those with multi-site subject agreement marking in position P1 as well as P8.

transitive configurations, along with a residue of unproductive ones. The examples in (4) show several verb stem formulas. Lexical morphemes are transcribed phonologically, the abbreviations SBJ and OBJ indicate agreement marker locations,⁷ and the material in parentheses identify the tense–mood–aspect class by showing the markers that the given stem requires in order to build present indicative, past indicative, and imperative forms (the language’s only three grammatical tense–mood categories):

(4) Sample Ket finite verb stems⁸

- a. SBJ⁸-OBJ^{4/3/1}-(i/il)-bɛd⁰‘SBJ makes OBJ’
 di⁸-b³-bɛd⁰ ‘I make it’
 di⁸-b³-il²-bɛd⁰ ‘I made it’
- b. SBJ⁶-k⁵-(a/ɔn)-dɛn⁰ ‘SBJ leaves’
 bɔ⁶-ɣ⁵-a⁴-tn⁰ ‘I go, I leave’
 bɔ⁶-ɣ⁵-ɔ⁴-n²-dɛn⁰ ‘I went, I left’
- c. SBJ⁸-aga⁷-k⁵-(s/il)-aq⁰ ‘SBJ goes to the forest’
 d⁸-aɣa⁷-k⁵-s⁴-aq⁰ ‘I go to the forest’
 d⁸-aɣa⁷-k⁵-l²-aq⁰ ‘I went to the forest’
- d. SBJ⁸-aga⁷-OBJ⁶-k⁵-(s/il)-aq⁰ ‘SBJ takes OBJ to the forest’
 d⁸-aɣa⁷-i⁶-k⁵-s⁴-aq⁰ ‘I take her to the forest’
 d⁸-aɣa⁷-i⁶-k⁵-l²-aq⁰ ‘I took her to the forest’
- e. SBJ⁸-igda⁷-OBJ⁶-k⁵-(s/il)-aq⁰ ‘SBJ takes OBJ to the river’
 da⁸-igda⁷-bɔ⁶-k⁵-s⁴-aq⁰ ‘she takes me to the river’
 da⁸-igda⁷-bɔ⁶-k⁵-l²-aq⁰ ‘she took me to the river’

The positionally distinct morpheme shapes that build Ket verb forms often resist individual glossing. The P₅ determiner consonants, for example, are often semantically opaque. The meaning of the Po base itself may no longer be divisible from the rest of the stem structure.

The oldest and semantically most basic core of the Ket verbal lexicon requires a verb root in the Po base position, as in the first two stem formulas above: *di⁸-b³-il²-bɛd⁰* ‘I made it,’ *bɔ⁶-ɣ⁵-ɔ⁴-n²-dɛn⁰* ‘I left.’ However, all productive derivational formulas in Ket now place the semantically and phonologically most salient lexical morpheme in the incorporate position P₇, near the verb’s leftmost edge. Stems of this sort were shown in (4c–e), where the original Po root **qod* ‘walk’ has eroded to *aq* and is now found

⁷ The animate-plural subject suffix slot P_{–1} is not shown, as its presence in the stem formula can normally be predicted from the presence of the subject person agreement slot in P₈ (and P₁).

⁸ The object marker labeled by triple superscript ^{4/3/1} in (4a) and other examples signifies object marking in P₄ for 3rd person animate-class agreement, P₃ for inanimate-class agreement, and P₁ for agreement with speech-act-participants. This arrangement is typical of all transitives that do not mark their objects in P₆.

in verbs meaning ‘take’ as well as ‘go.’ The phonologically and semantically most salient portions of these verb forms are the incorporated directionals *aga* ‘from the river up into the forest’ and *igda* ‘from the forest to the river.’ Over the past few millennia, the Yeniseian verb seems to have undergone gradual realignment to a *de facto* suffixing structure in which the most salient lexical material occupies the verb’s initial syllable, probably in typological accommodation to the surrounding suffixing languages as influenced through multilingualism and exchange of marriage partners (Vajda 2009). In complex verbs with salient incorporates, the affixes in positions P6–P1 have come more to resemble suffixes.

Another incorporate-dominant verb is shown in (5), where the P7 infinitive *bagdeŋ* ‘to drag’ serves as the verb’s semantic head. The P5 determiner is the semantically opaque consonant *k*, which labializes the preceding object marker; other *k*-determiners, presumably of different origins, do not labialize the preceding vowel. The base position Po contains *bed*, a form that possibly originated through bleaching of the verb root *bēd* ‘make’ that appeared in examples (4a,b) examined earlier. In (5), however, the Po morpheme *bēd* simply helps to convey repetitive action and could almost be regarded as a derivational affix though it occupies the original verb root position:

- (5) SBJ⁸-bagdeŋ-Obj⁶-k⁵-(a/ɔl)-bed⁰ ‘SBJ drags OBJ repeatedly’
 da⁶-bagdeŋ⁷-bo⁶-γ⁵-a⁴-bed⁰ ‘she drags me repeatedly’
 da⁶-bagdeŋ⁷-bo⁶-γ⁵-ɔ⁴-l²-bed⁰ ‘she dragged me repeatedly’

Example (5) illustrates a highly productive stem creation pattern. Virtually any infinitive can be inserted into this formula to produce an iterative verb with the same lexical meaning as the infinitive itself.

Other productive patterns are shown in (6), where any infinitive can be inserted into formulas containing the P5 determiner *q*, probably originally meaning ‘movement inside,’ but here signaling the beginning of action. The resultant stem means: ‘SBJ begins action once’ (6a). The iterative inceptive in (6b) is built by substituting conjugation class *a/ol* and Po base *-da*:

- (6) a. SBJ⁸-bagdeŋ-q⁵-Obj^{4/3/1}-(i/in)-t⁰ ‘SBJ starts dragging OBJ once’
 da⁶-bagdeŋ⁷-q⁵-i⁴-di¹-t⁰ ‘she starts dragging me once’
 da⁶-bagdeŋ⁷-q⁵-i⁴-n²-di¹-t⁰ ‘she started dragging me once’
 b. SBJ⁸-bagdeŋ-q⁵-Obj^{4/3/1}-(a/ɔl)-da⁰ ‘SBJ starts dragging OBJ (iter.)’
 da⁶-bagdeŋ⁷-q⁵-a⁴-d¹-da⁰ ‘she starts dragging me repeatedly’
 da⁶-bagdeŋ⁷-q⁵-ɔ⁴-l²-di¹-da⁰ ‘she started dragging me repeatedly’

Active intransitive inceptives are productively made according to the formula exemplified by (7), which contains the incorporated infinitive *iʹl* ‘to sing, singing, song.’ Once again, any infinitive can be substituted to produce an intransitive verb expressing inception of an activity:

- (7) il^7 - SBJ^6 - k^5 -(a/ω)- qan^0 ‘SBJ starts singing’
 il^7 - ba^6 - γ^5 - a^4 - κan^0 ‘I start singing’
 il^7 - ba^6 - γ^5 - ω^4 - $\kappa \omega n^0$ ‘I started singing’

Notice that the stem in (7) contains non-labializing *k*-determiner. Productive patterns involving adjective or noun incorporation require a *t*-determiner (also semantically opaque):

- (8) a. $NOUN^7$ - SBJ^6 - t^5 -($a/\omega n$)- aq^0 ‘SBJ becomes NOUN’
 $\kappa \epsilon t^7$ - ba^6 - t^5 - ω^4 - n^2 - aq^0 ‘I became a man’ < $\kappa \epsilon$ ’*t* ‘person’, aq ‘become’
 $d \epsilon \eta^7$ - $d \Lambda \eta^6$ - t^5 - ω^4 - n^2 - aq^0 ‘we became men’ < $d \epsilon$ ’*η* ‘people’, aq ‘become’
- b. SBJ^8 - ADJ^7 - t^5 - $OBJ^{4/3/1}$ -($a/\omega l$)- sin^0 ‘SBJ turns OBJ into ADJ’
 da^8 - qa^7 - t^5 - a^4 - b^3 - sin^0 ‘she makes it big’ < qa ’*big*, sin ‘cause to be’
 k^8 - $ugde^7$ - t^5 - ω^4 - l^2 - sin^0 ‘you lengthened him (masc.-class rabbit coat)’
< $ugde$ ’*long*, sin ‘cause to be’
- c. SBJ^8 - $NOUN^7$ - OBJ^6 - h^5 -($a/\omega l$)- ted^0 ‘SBJ hits OBJ with NOUN’
 da_8 - $ulgit^7$ - ba^6 - γ^5 - a^4 - $t \epsilon d^0$ ‘she elbows me’ < $ulgit$ ’*elbow*, $t \epsilon d$ ‘hit’
 da^8 - $ulgit^7$ - ba^6 - γ^5 - ω^4 - l^2 - $t \epsilon d^0$ ‘she elbowed me’
 d^8 - $d \omega n^7$ - a^6 - γ^5 - a^4 - $t \epsilon d^0$ ‘I stab him’ < $d \omega$ ’*n* ‘knife’, $t \epsilon d$ ‘hit’
 d^8 - $d \omega n^7$ - a^6 - γ^5 - ω^4 - l^2 - $t \epsilon d^0$ ‘I stabbed him’
- d. SBJ^8 - $NOUN^7$ -(s/il)- $b \epsilon d^0$ ‘SBJ makes NOUN’
 da^8 - qus^7 - si^4 - $b \epsilon d^0$ ‘she makes a tent’ < qu ’*s* ‘tent’, $b \epsilon d$ ‘make’
 da^8 - qus^7 - li^2 - $b \epsilon d^0$ ‘she made a tent’
 d^8 - qun^7 - si^4 - $b \epsilon d^0$ - n^{-1} ‘they make tents’ < qu ’*η* ‘tent’, $b \epsilon d$ ‘make’
 d^8 - qun^7 - li^2 - $b \epsilon d^0$ - n^{-1} ‘they made tents’

Examples (5)–(8) show several of the most important patterns from among the few dozen productive templatic derivational formulas of Modern Ket, where lexical morpheme shapes, conjugation classes, and agreement configurations interact to create new finite verb stems. Similar processes operated in Yugh and Kott.

28.4 INFINITIVES

Examining the Yeniseian infinitive (a type of action nominal) can serve as a sort of conclusion, as this class of words displays a unique combination of morphological traits linking it with both nouns and finite verbs, as well as with modifiers. Infinitives are historically derived from the finite verb stem template according to an ancient pattern that shows unmistakable homologies with Na-Dene (Vajda 2011: 60–3). Basically, in Yeniseian the lexical morphemes in P7, P5, and Po combine to form the corresponding

infinitive, with omission of any morphological material that would normally occupy the intervening positions in the corresponding finite verb. In the case of finite verb stems containing no lexical morphemes other than the root, infinitive derivation originally required an epenthetic element called a peg prefix, probably in keeping with the fact that finite-verb forms were (originally) never root-initial. The peg prefix appears regularly in Kott infinitives made from simple verb stems: *ši-ten* ‘to lie down,’ *ši-puk* ‘to pull, drag,’ *ši-kit* ‘to rub,’ *ši-ti* ‘to beat,’ *ši-pei* ‘to hunt,’ etc. Infinitives that inherit an incorporate or determiner morpheme from the corresponding finite verb lack this peg prefix: Kott *ax-kit* ‘to plane wood,’ *tʰa-pei* ‘to pursue.’ Ket infinitives have almost completely lost the prefix (*kɪt* ‘to rub,’ *tār* ‘to beat’), but often contain an infinitive suffix that varies inexplicably between the allomorphs *-ɛj*, *-ɛŋ*, and reduction to falling tone in the preceding vowel: *bɛr-ɛŋ* ~ *bèr* ‘to make,’ *si-bagd-ɛŋ* ~ *bagd-ɛŋ* ‘to pull, drag.’ The infinitive suffix survives in Kott-Assan only sporadically (e.g. Assan *taŋ-n* ‘to see,’ but Kott *tʰaŋ* ‘to see’). Proto-Yeniseian infinitive derivation thus seems to have followed the general pattern of: (1) finite verb incorporate–determiner–root + infinitive suffix **-əŋ*; or in the case of simple stems: (2) peg prefix **si-* + finite verb root + infinitive suffix **-əŋ*.

Though based on the finite verb template, Yeniseian infinitives are a type of modifier in the sense that, like adjectives, they can be preposed to nouns: *bagdɛŋ kɛt* ‘dragging person,’ ‘a person who drags or is dragged.’ Also like adjectives, they take predicate concord suffixes such as 1SG *-di*, in which case they express the ability to perform an action: *a t il-di* ‘I can sing’ (Krejnovich 1968: 26). Infinitives resemble nouns, however, in being able to take possessive prefixes: *b-daŋ* ‘my living, my life,’ *da-daŋ* ‘her living, her life,’ *k-iŋ* ‘your singing, your song,’ *d-iŋ* ‘his singing, his song,’ etc. Possessed infinitives associated with events that logically involve both actor and undergoer such as *d-bagdɛŋ* ‘his dragging’ are ambiguous between the meanings ‘his being dragged’ and ‘his dragging (of something or someone else).’ Unlike true nouns, infinitives are regularly nominalized with *-s* to create deverbal nouns. Nominalized infinitives such as *bagdɛŋs*, like their possessed counterparts, are ambiguous between meanings such as ‘someone who drags’ and ‘someone or something that gets dragged.’ The Kott forms *ši-čer-še* ‘scraper’ (< *ši-čer* ‘to scrape’) and *ši-čal-še* ‘file’ (< *ši-čal* ‘to sharpen’) show that the nominalization of infinitives, like infinitive derivation itself, follows an ancient, inherited pattern.

CHAPTER 29

MON-KHMER

MARK J. ALVES

29.1 OVERVIEW

THIS chapter presents an overview of derivational word-formation processes¹ among Mon-Khmer languages, of which the main strategies involve the use of prefixes and infixes, compounding, and a type of reduplication involving phonological alternations. While some core, largely fossilized processes, such as causative prefixes and nominalizing infixes, are nearly universal in this language group and even reconstructable to the proto level, there are many other semantico-syntactic functions of Mon-Khmer morphology spread out less evenly in the region.

29.2 MON-KHMER HISTORICAL CONTEXT, PREVIOUS RESEARCH, AND MAIN ISSUES

The term Mon-Khmer has traditionally referred to a sub-branch of Austroasiatic, in which Munda is considered a sister sub-branch, though recently, this relationship has been questioned, with the hypothetical possibility that Munda developed from Mon-Khmer (Diffloth 2005, Sidwell 2008). The identification of the Mon-Khmer language

¹ As this chapter is focused on derivational morphology, inflectional morphology is not described. This includes subject-marking on verbs (Aslian), case-marking on pronouns (Katuic and Vietic), case-marking on nouns (Mangic), plural nouns (Aslian) and plural adjectives/stative verbs (Aslian and Katuic), and aspect (Aslian, Bahnaric, Katuic, and Vietic), among other types of morphology that have grammatical functions. For more information on grammatical morphology among Mon-Khmer languages, see “Morphological Functions among Mon-Khmer Languages: Beyond the Basics” (Alves forthcoming). Related materials include presentation handouts posted online for the 2012 Max Planck workshop on Mainland Southeast Asian typology and the 2013 Annual Conference of the Society of Southeast Asian Linguistics.

family goes back to the mid- to late-1800s (Gage 1985: 495). At the beginning of the twentieth century, Schmidt (1906) first posited the Austroasiatic and Austronesian language families that he grouped under Austric, a kind of supra-language family, though this latter point is still debated (Reid 1999). The typological differences between Munda and Mon-Khmer are massive (Donegan and Stampe 1983, Donegan 1993). Mon-Khmer syntax is mostly analytic and its morphology is largely isolating (e.g. no morphological marking of plurality on nouns or morphological aspect on verbs in most, but not all, Mon-Khmer languages). Affixation in Mon-Khmer is, with a small number of exceptions, largely derivational, with a large quantity of fossilized prefixes and infixes, but no suffixes. Munda languages, in contrast, have agglutinative and polysynthetic morphology, including both inflectional and derivational affixes, including prefixes, infixes, as well as suffixes. As Munda is so typologically different, only Mon-Khmer is dealt with in this study. For an overview of Munda word formation, see Pinnow 1966, Zide and Anderson 2001, and Anderson 2004 and 2007.

Mon-Khmer consists of some dozen sub-branches (with additional hypothesized sub-branches) having over 160 languages throughout mainland Southeast Asia and neighboring parts of eastern India and southern China. Common basic vocabulary and shared causative prefixes and nominalizing infixes have been identified in all these languages and demonstrate genetic affiliation. No sub-branching of these groups has been solidly established. Vietnamese, which is typologically distinct from the rest of the family at least in part due to massive language contact with Chinese, has the largest population of speakers, while speakers of some Mon-Khmer languages, such as related Vietic languages Ruc and Sach, number only in the hundreds. Most names of the sub-branches come from relatively larger groups of speakers who sometimes number in the dozens or even hundreds of thousands of speakers, but Khmer is the only Mon-Khmer language other than Vietnamese with more than ten million speakers. The vast majority of Mon-Khmer languages are languages of ethnic minorities, many of whom are culturally and linguistically in danger of being lost to larger linguistic communities.

Table 29.1 lists the sub-branches of Mon-Khmer referred to in this study with about thirty key publications that describe morphology among the languages. The publications are mostly in English, though for languages spoken in Vietnam, many are Vietnamese language publications. While the list is not short and consists of numerous of available major works, it is not exhaustive. Moreover, in many cases, information on Mon-Khmer languages consists of single articles, many of which provide little or no description of morphology. More often, no linguistic descriptions exist at all, and the languages are merely items on lists of languages. Considering the large number of languages in this family, this study is still just a sample.

As the descriptions vary in the amount of detail offered, there are cases in which it is difficult to determine with certainty whether posited affixes are actually affixes or clitics instead. In most categories in this study, the phonological material is clearly lexicalized. For instance, in addition to not being separable, there are inconsistent phonological patterns, and the elements do not attach freely to all categories of words. There are, however, instances in which the posited affixes have existing cognate sources, in which case, it could be argued that these are simply phonologically

Table 29.1 Core references on Mon–Khmer morphology

Sub-branches	Languages and primary sources
Aslian (Peninsular Malaysia)	Entire branch (Matisoff 2003); Jah–Hut (Diffloth 1979); Jahai (Schebesta 1926, 1931, Burenhult 2002); Kentakbong (Omar 1976); Semelai (Kruspe 2004); Semai/Sengoi (Diffloth 1976a, 1976b, Means and Means 1986); Temiar (Benjamin 1976, 2011, Means 1998)
Bahnaric (Central Vietnam)	Bahnar (Banker 1964a, 1964b); Chrau (Thomas 1969, Thomas 1971); Halang (Cooper 1966); Jeh (Cohen 1966, Gradin 1976); Koho (Hoàng Tuệ et. al. 1986, Tạ 1998, Tạ Văn Thông 1998); Mnong (Dinh Le Thu 2007); Rengao (Gregerson 1979); Sedang (Smith 1969); Sre (Manley 1972)
Katuic (Central Vietnam)	Entire branch (Hoàng Văn Hành 1993); Bru (Nguyễn Văn Tài et al. 1984; Vương, Hữu Lễ 1997, Hoàng Văn Ma and Tạ Văn Thông 1998); Katu (Costello 1966, 1998, 2001, Nguyễn Hữu Hoành and Nguyễn Văn Lợi 1998; Nguyễn Hữu Hoành 1998); Ngeq (Smith 1973); Pacoh (S. Watson 1964, 1966; R. Watson 1966a, 1966b, 1976, 2011, Nguyễn Văn Lợi et al. 1986; Alves 2004)
Khasic (Northeastern India)	Khasi (Gurdon 1914; Nagaraja 1985)
Khmeric (Cambodia)	Khmer (Karnchana 1978, Jenner and Pou 1982, Thomas 1987, Thomas 1990)
Khmuic	Khmu (Svantesson 1983, Premssirat 1987); Mal (Filbeck 1996); Mlabri (Rischel 1995)
Mangic/Pakanic	Bugan (Li 1996, Li 2005); Mang (Nguyễn Văn Lợi et al. 2008)
Monic (Thailand and Burma)	Mon (C. Bauer 1982, 1989; Jenny 2003, 2005)
Nicobarese (Nicobar Islands of India)	Nancowry Nicobarese (Man 1889; Radakhrisnan 1976 and 1981)
Palaungic (Burma and China)	Palaung (Milne 1921, Shorto 1963); Riāng–Liang (Shorto 1963)
Vietic (Vietnam)	Kri (Enfield and Diffloth 2009); Muong reduplication (Hoàng Văn Hành 1998); Ruc (Nguyễn Văn Lợi 1993); Vietnamese (Thompson 1965, Nguyễn Tài Cẩn 1996, Nguyễn Đình Hoà 1966); Vietnamese reduplication (Hoàng Văn Hành 1985, Viện Ngôn Ngữ Học (Institute of Linguistics, Hanoi) 1995, Vu 2007);
Pearic (Cambodia and Thailand)	Pearic (Headley 1977; Suwilai 2011)
General issues in Mon–Khmer	Affixes (C. Bauer 1990, Alves 2001a, Sidwell 2008); Pronouns (Pinnow 1965); Typology (Donegan and Stampe 1983, Donegan 1993)

reduced words or clitics. R. Watson (2011) provides a description of the Katuic language Pacoh showing how elements which appear affix-like are actually clitics, and similar phenomena are seen among other Mon–Khmer languages, most notably among subject-position preverbal elements, an inflectional category not part of this

study. Where relevant, these concerns are recognized in the chapter, but whether a technical definition includes or excludes these from the category of affixes is something that is beyond the scope of this chapter. Nevertheless, overall, the data sources for this study cover what are, for the most part, historical remnants of a more productive affix system, with some more productive morphological material, though these are the exception and not the rule.

The most common categories of word-formation in Mon-Khmer languages include those in Table 29.2. The degree of productivity of the morphology varies, but most often, (a) affixation is largely fossilized with very limited productive capacity, (b) reduplication is somewhat more productive with a good degree of creative capacity and a significant range of phonological variety (as discussed in subsequent sections), and (c) compounding is a fairly open process, as expected. While some types are common throughout language family, such as deriving nouns from verbs, there is significant variation in what can be derived from what. For instance, in various Mon-Khmer languages, a causative verb can be derived from a verb, a noun, or other lexical subcategories. All of these types are seen in all available descriptions of these languages, with the exception of Mang (Mangic), which has no infixes, and Vietnamese, a robustly isolating language with phonaesthemes or reconstructable affixation at best, though in reduplication, Vietnamese is quite rich.

These three main categories are described in following sub-sections. Each section starts with a brief description of the phonological aspects followed by descriptions of the various semantic and syntactic subcategories.

Table 29.2 Common derivational processes in Mon-Khmer

-
1. Affixation
 - a. Prefixes and infixes deriving verbs (causative verbs, stative verbs/adjectives,* reciprocal, etc.)
 - b. Prefixes and infixes deriving nouns (instrumental, resultative, locative, etc.)
 - c. Prefixes to derive personal and demonstrative pronouns
 2. Reduplication
 - a. Complete reduplication (generalizing adverbs, degree of intensity, etc.)
 - b. Alternating reduplication (a wide range of semantic types of nouns, active verbs, stative verbs/ adjectives, and adverbs)
 3. Compounding
 - a. Left-headed
 - b. Dual-headed
 - c. Pseudo-compounding
-

Note: * The distinction between stative verbs and adjectives is not always syntactically distinct among Southeast Asian languages. They are essentially the same for the purposes of this study.

29.3 AFFIXATIONAL MORPHOLOGY IN MON-KHMER

This section describes affixes in Mon-Khmer, starting with the issue of phonological shape, followed by the various semantic and semantico-syntactic categories of these affixes.

29.3.1 Phonological Issues

In many Mon-Khmer languages, in main syllables, CV(C) is the standard basic syllable shape, initial consonant clusters are a good deal less common (varying according to language), and initial stops in presyllables are often voiceless. Consonant clusters consist most often of stops plus liquids (e.g. /pl/, /tr/, etc.), and vowel systems can be differentiated by both length and phonation (e.g. clear vs. breathy), leading to rather large vowel inventories of two dozen or more vowel phonemes. Bisyllabic words take stress on the final syllables so much that unstressed presyllables are phonologically reduced. The vowels of presyllables are often restricted to schwa or another central, neutral vowel, or even no full vowel, a phenomenon known as “sesquisyllabism” (Matisoff 1973). In some sub-branches, such as Aslian and Katuic, other vowels, such as /i/, /a/, and /u/, occur in presyllables, but these are the exception. Presyllables can even be nasals, often assimilating to following segments. Assimilation to the place of articulation (e.g. labials, alveolars, palatals, and velars) is not uncommon, as shown by the following samples of the causative prefix having the general shape /pɔN-/ in Khmer.

- (1) Khmer (Khmeric): /kʰos/ ‘wrong’ → /bɔŋ-khos/ ‘to cause a mistake’; /co:l/ ‘to enter’ → /bɔŋ-co:l/ ‘to cause to enter’; /to:c/ ‘small’ → /bɔn-to:c/ ‘to diminish in power’ (Karnchana 1978: 193)²

29.3.2 Categories of Affixes and Productivity

The method applied in this study was to mark a category positive for a feature even if only one language in that branch has that feature, though when a sub-branch does

² In this chapter, IPA transcriptions of the original sources are made, with an attempt to keep the overall phonological representation. Syllabification is not indicated except where morphological boundaries are indicated with hyphens. Vowel length is a complex issue in Mon-Khmer languages, and as authors use a variety of means to indicate vowel length, in this study, only long vowels in closed syllables are marked with a colon. Original sources are consistently indicated so that questions about the phonological forms can be checked.

Table 29.3 Types of affixation in Mon-Khmer languages

Sub-branch categories	KS	MG	PL	KU	AS	NC	MN	KR	BN	KT	VT
1. Causative	+	+	+	+	+	+	+	+	+	+	+
2. Nominalizing	+	+	+	+	+	+	+	+	+	+	+
3. Reciprocal	–	–	+	+	–	–	–	+	+	+	–
4. Stativizing	+	–	–	+	+	–	–	–	+	+	–
5. Pronominal	+	–	+	+	–	–	–	–	+	+	–
6. Completive	–	–	–	+	+	–	–	–	+	+	–
7. Verbalizing	–	+	–	+	+	–	–	–	+	–	–
8. Desiderative	+	–	–	–	+	–	–	–	–	+	–
Total number	5	3	4	7	6	2	2	3	7	7	2

Note: the following abbreviations are used for the various sub-branches: AS-Aslian, BN-Bahnaric, KR-Khmeric, KS-Khasic, KT-Katuic, KU-Khmuic, MG-Mangic, MN-Monic, NC-Nicobarese, PL-Palaungic, VN-Vietnamese, VT-Vietic.

have several languages, typically, more than one language has that feature.³ More languages and sub-branches probably have these features, but the available literature does not clearly indicate so, and thus, this is a preliminary finding. A distinction has been made in this study between those morphological categories seen in three or more sub-branches, as in Table 29.3, and those in only two sub-branches, which are treated collectively in Section 29.3.3. It is important to note that in Table 29.3, VT (Vetic) refers not to Vietnamese and the closely related Muong languages as they do not have affixes, but rather to groups of speakers of other Vetic languages which are polysyllabic/sesquisyllabic like other typical Mon-Khmer languages. Total numbers of functions in each branch are listed in this and other tables in this chapter.

The degree of productivity in many instances is generally limited, with a few exceptions. Most often, existing Mon-Khmer affixation merely suggests earlier word-formation productivity. For instance, in Khmer, while Table 29.3 would suggest a very limited range of affixation, Jenner and Pou (1982: xxxiii–lviii) describe a few dozen distinct semantic functions of affixes, such as potentials, intensives, locatives, and so on, although admittedly in almost exclusively fossilized/lexicalized forms. Throughout Mon-Khmer, causative prefixes and nominalizing infixes are most often lexicalized but have also merged with other morphological material to create newer semantically

³ As mentioned, dozens of these languages have very limited or simply no complete descriptions published. This is a recognized weakness in this study, but it cannot be avoided without additional data, which is not likely to be available in the near future. It is best for now to publish and let future investigations either support or refute the claims in it when it is possible to do so.

Table 29.4 Causative affixes in Mon–Khmer

Categories	Sample allomorphs
Single initial consonants	/p-/ , /p ^h -/
Consonant clusters	/pN-/ , /pr-/ , etc.
Presyllables	/pə-/ , /pa-/ , /pi-/ , /pV-/ , /pən-/ , /pin-/ , /pik-/ , /par-/ , /bə-/ , /bVŋ-/ , /bVr-/ , /tə-/ , /ta-/ , /tən-/ , /ʔa-/ , /ka-/ , /sa-/ , /ha-/ , /mə-/ , etc.
Infixes	/-m-/ , /-um-/ , etc.

complex forms. Reciprocal affixes are somewhat more productive in some languages. The categories in Table 29.3 are described in the following subsections.

29.3.2.1 *Causative Prefixes*

Causative prefixes, which often derive causative transitive verbs from intransitive verbs, are universal in Austroasiatic, as they are in Mon–Khmer as well as Munda languages.⁴ These prefixes tend to have the general phonological shape of /pV-/ or /pVN-/ , though there is considerable diversity, as shown in Table 29.4.

These prefixes always derive verbs, but the source roots, which most often are active verbs, can also be stative verbs and even nouns, as in examples (6), (9), and (10) below. The syntactic structures of complements vary significantly among the sub-branches and languages, ranging from basic transitive verbs to verbs requiring multiple complements, locations, and even lower clauses (Alves 2001a).

- (2) Khasi (Khasic): /ba:m/ ‘to eat’ → /pin-ba:m/ ‘to feed’ (Nagaraja 1985: 27)
- (3) Khmu (Khmuic): /kà:j/ ‘to come’ → /p-kà:j/ ‘to cause to come’ (Premssirat 1987: 26)
- (4) Mang (Mangic): /əit⁶/ ‘to die’ → /pa-əit⁶/ ‘to kill’ (Nguyễn Văn Lợi, Nguyễn Hữu Hoàn, and Tạ Văn Thông 2008: 141)
- (5) Mlabri (Khmuic): /ʔem/ ‘to sleep’ → /pa-ʔem/ ‘to cause to fall asleep’ (Rischel 1995: 304)
- (6) Nancowry (Nicobarese): /káh/ ‘to know’ → /ha-káh/ ‘to cause to know’ (Radhakrishnan 1981: 55)
- (7) Palaung (Palaungic): /ʔa:h/ ‘to be large’ → /pən-ʔah/ ‘to make large’ (Milne 1921: 74)
- (8) Rengao (Bahnaric): /hlat/ ‘to die’ → /ba-hlat/ ‘to kill’ (Gregerson 1979: 94)
- (9) Riang-Liang (Palaungic): /lɛ/ ‘to go out’ → /p-lɛ/ ‘to drive out’ (Shorto 1963: 53)
- (10) Sedang (Bahnaric): /dan/ ‘be erect’ → /mə-dan/ ‘to raise’ (Smith 1969: 113)

⁴ Even in Vietnamese, there are hints of earlier causative prefixes, such as *chết* ‘to die’ versus *giết* ‘to kill’, showing the possibility of an earlier prefix on the latter item.

Table 29.5 Nominalizing affixes in Mon-Khmer

Types	Allomorphs
Prefixes	/bə-/ , /m-/ , /ma-/ , /m-/ , /sN-/ , /sr-/ , /tr-/ , etc.
Infixes	/-n-/ , /-an-/ , /-ən-/ , /-əm-/ , /-mn-/ , /-rm-/ , /-nr-/ , /-m-/ , /-n-/ , /-nt-/ , /-əd-/ , /-əŋ-/ , /-r-/ , /-rn-/ , /-mp-/ , /-ərn-/ , /-arn-/ , etc.

- (11) Semelai (Aslian): /kasut/ ‘shoes’ → /p-kasut/ ‘to put shoes on someone’ (Kruspe 2004: 83)

Bisyllabic prefixes are rare in Mon-Khmer, with exceptions in the Katuic and Nicobarese sub-branches, in which they do appear typically as a result of combinations of affixes.

- (12) Katu (Katuic): /sarum/ ‘to fall’ → /tapa-sarum/ ‘to cause to fall’ (Costello 1966: 64)
- (13) Nancowry (Nicobarese): /kīa/ ‘to close an eye’ → /ha-kīa/ ‘to aim at’ → /ma-ha-kīa/ ‘one who aims at’ (Radhakrishnan 1981: 57)

29.3.2.2 *Nominalizing Affixes*

A universal means of nominalization in Mon-Khmer languages is nasal infixes deriving nouns from verbs, making this a feature of proto-Mon-Khmer with additional semantic distinctions and corresponding phonological reconstructions at that level (Sidwell 2008). These forms often show assimilation to place of articulation, as shown by examples in the Aslian language Sengoi. The nominalizing infix is /-ən-/ before vowels, /-ŋ-/ before non-labial consonants, and /-m-/ before labial consonants (Table 29.5).

- (14) Sengoi (Aslian): /bet/ ‘to sleep’ → /b-ən-ət/ ‘sleep’; /besat/ ‘to instruct’ → /bəŋ-sat/ ‘instruction’; /pəp/ ‘to touch’ → /pəm-pəp/ ‘touch’ (Means and Means 1986: 8)

In many other Mon-Khmer languages, allomorphs of these fossilized elements have no particular phonological pattern. They occur as infixes in most Mon-Khmer languages, though in a few languages, they are prefixes, or even completely fused phonetic elements, as seen in Vietnamese deictics in Table 29.6.

Table 29.6 Demonstratives in Vietnamese

	Unspecified	Close to speaker	Remote
Place	ɗəw ¹ (đâu)	ɗaj ¹ (đây)	ɗəj ³ (đấy)
Reference	na:w ² (nào)	naj ² (này)	no ⁶ (nó)
Proportion	ba:w ¹ (bao)	bəj ¹ (bây)	bəj ³ (bấy)

Furthermore, in some languages, such as Khasi in (20) below, other non-cognate prefixes are used to derive nouns. One example of this variety is in Khmu, which has been shown to have at least four forms of both prefixes (/rn-/ , /sN-/ , /sr-/ , /tr-/) and infixes (/mn-/ , /-m-/ , /-n-/ , /-rn-/) (Svantesson 1983).

(15) Khmer (Khmeric): /sɗap/ 'to listen' → /s-ɔn-dap/ 'teaching' (Karnchana 1978: 196)

(16) Ruc (Vietic): /tʰut³/ 'to stop up' → /t-an-ut³/ 'stopper' (Nguyễn Văn Lợi 1993: 83)

(17) Nancowry (Nicobarese): /tún/ 'to smell' → /t-am-ún/ 'one who smells' (Radhakrishnan 1981: 56)

Semantically, they consist of several noun sub-classes. In the following example in Jah-Hut ((18)), an Aslian language, the source verb meaning 'to eat' is the root from which a gerundive, agentive, and resultative are derived. Such a complete set of derived forms is rather exceptional, but it is suggestive of the kinds of patterns in the language family.

(18) Jah-Hut (Aslian): /caʔ/ 'to eat'; /nʔ-caʔ/ 'the act of eating'; /mʔ-caʔ/ 'one who eats too much'; /pɲ-caʔ/ 'food' (Diffloth 1979: 98)

Nothing in the literature suggests any geographic concentrations of types of derived nominals. In the following samples, the several semantic classes of derived nouns are described and exemplified, including general nominal, agentive, resultative, instrumental, and locative nouns. General nominals are common, as expected (Table 29.7).

Table 29.7 Specialized semantico-syntactic functions of affixation in Mon-Khmer

Sub-branch categories	KS	MG	PL	KU	AS	NC	MN	KR	BN	KT	VT
1. Negation	-	-	-	-	-	-	+	+	+	+	-
2. Existential/locative	-	-	-	-	-	-	-	-	-	+	+
3. Involuntary	-	-	-	-	-	-	-	-	+	+	-
4. Number of days/months	-	-	-	+	-	-	-	-	-	+	-
5. Stative verbs with plural subjects	-	-	+	-	-	-	-	-	-	+	-
6. Pretence	-	-	+	-	-	-	-	-	-	+	-
Total number	0	0	2	1	0	0	1	1	2	6	1

Note: the following abbreviations are used for the various sub-branches: AS-Aslian, BN-Bahnaric, KR-Khmeric, KS-Khasic, KT-Katuic, KU-Khmuic, MG-Mangic, MN-Monic, NC-Nicobarese, PL-Palaungic, VN-Vietnamese, VT-Vietic.

- (19) Katu (Katuic): /pəŋ/ ‘to wear on head’ → /p-an-əŋ/ ‘hat’ (Nguyễn Hữu Hoàn 1995: 236)
 (20) Khasi (Khasic): /trej/ ‘to work’ → /(ka) jɪŋ-trej/ ‘(the) working’ (Nagaraja 1985: 10)
 (21) Palaung (Palaungic): /pɤm/ ‘to kill’ → /pən-pɤm/ ‘the killing/execution’ (Milne 1921: 73)

Agentives, on the other hand, refer specifically to agents of actions.

- (22) Khasi (Khasic): /trej/ ‘to work’ → /(u/ka) nɔŋ-trej/ ‘(the male/female) worker’ (Nagaraja 1985: 9)
 (23) Nancowry (Nicobarese): /tuŋ/ ‘to smell’ → /t-an-uŋ / ‘one who smells’ (Radhakrishnan 1981: 56)

Some nominalized forms refer to the resulting objects of the actions.

- (24) Bahnar (Bahnaric): /tar/ ‘to weave’ → /t-ən-ar/ ‘woven bamboo’ (Banker 1964a: 101)
 (25) Bru (Katuic): /paj/ ‘to speak’ → /p-arn-aj/ ‘speech’ (Nguyễn Văn Tài, Đoàn Văn Phúc, and Phan Xuân Thành 1984: 37)
 (26) Khmer (Khmeric): /bok/ ‘to pound (rice)’ → /ʔam-bok/ ‘roasted, pounded glutinous rice’ (Jenner and Pou 1982: xlvi)

These also derive instrumental nouns which are semantically the means by which the actions occur.

- (27) Katu (Katuic): /dum/ ‘to lie in a blanket’ → /d-ərŋ-um/ ‘wing’ (Nguyễn Hữu Hoàn 1995: 237)
 (28) Khmu (Khmuic): /criəŋ/ ‘to comb’ → /c-nt-riəŋ/ ‘a comb’ (Premsrirat 1987: 27)
 (29) Koho (Bahnaric): /kɔr/ ‘to scrape’ → /k-ən-ɔr/ ‘a scraper’ (Hoàng Tuệ, Lý Toàn Thắng, Tạ Văn Thông, et. al. 1986: 55)

Locative nouns are also derived from these affixes.

- (30) Bahnar (Bahnaric): /puŋ/ ‘to wallow’ → /p-ən-uŋ/ ‘wallowing hole’ (Banker 1964a: 102)
 (31) Khmer (Khmeric): /tom/ ‘to perch’ → /t-ran-om/ ‘a perch for chickens’ (Dorothy Thomas 1990: 88)
 (32) Kri (Vietic): /ko:ʔ/ ‘to live’ → /k-rn-o:ʔ/ ‘a house’ (Enfield and Diffloth 2009: 47)

29.3.2.3 *Reciprocal Prefixes and Infixes*

Reciprocal affixes are common in Mon-Khmer languages, though not documented in all sub-branches. In some languages, these affixes are infixes or infixes in which two affixes have phonologically merged into one (C. Bauer 1990), while in other languages,

they are strictly prefixes, but regardless, there is a basic tendency for an /-r-/ element in either type. They are highly productive in some Mon-Khmer languages, such as Bru (Katuic) and Bahnar and Chrau (Bahnaric), attaching very freely to active verbs.

- (33) Chrau (Bahnaric): /pəm/ 'to hit' → /tam-pəm/ 'to fight' (Thomas 1971: 154)
- (34) Khmer (Khmeric): /cum/ 'to surround' → /pra-cum/ 'to join together' (Jenner and Pou 1982: 85)
- (35) Khmu (Khmuic): /pók/ 'to bite' → /trópók/ 'to bite each other' (Svantesson 1983: 39)
- (36) Palaung (Palaungic): /ʔɛ:h/ 'to scold' → /kəʔ-ʔɛ:h/ 'to abuse each other' (Milne 1921: 52)
- (37) Rengao (Bahnaric): /ʔwaj/ 'to reside' → /ta-ʔwaj/ 'to live together' (Gregerson 1979: 108)

Bru is an example of the variation according to root shape. Bru /ra-/ occurs on monosyllabic roots, while /-ar-/ occurs in the presyllables of bisyllabic roots.

- (38) Bru (Katuic): /cɛʔ/ 'to change' → /ra-cɛʔ/ 'to exchange with each other'; /sabaw/ 'to call' → /sa-r-baw/ 'to call each other' (Hoàng Văn Ma and Tạ Văn Thông 1998: 82)

29.3.2.4 *Stative Prefixes*

Many Mon-Khmer languages have prefixes to derive stative verbs with resultative adjectival and even passive-like meanings. Sidwell (2008: 14) has tentatively posited a proto-form in Mon-Khmer of *h-/hN-, though there is substantial variety in phonological form among the languages. For instance, Khmu has a rich array of allomorphs in this category: /m-/ , /n-/ , /ŋ-/ , /nt-/ , /ŋk-/ , /mp-/ , and /tr-/ (Premsrirat 1987: 26). Even Ruc (Vietic) has historical traces of a prefix on a number of adjectives (Nguyễn Văn Lợi 1993: 88), though the roots have largely been lost, leaving bisyllabic words with tentative prefixes.

- (39) Chrau (Bahnaric): /rih/ 'to tear' → /ta-rih/ 'ripped' (Thomas 1971: 153)
- (40) Katu (Katuic): /jah/ 'to break' → /ha-jah/ 'to be broken' (Nguyễn Hữu Hoàn 1995: 241)
- (41) Khasi (Khasic): /b^ha/ 'goodness' → /ba-b^ha/ 'good' (Gurdon 1914: 109)
- (42) Khmu (Khmuic): /lāk/ 'to split' → /hⁿ-lāk/ 'split' (Svantesson 1983: 71)
- (43) Pacoh (Katuic): /hɛ:ʔ/ 'to tear' → /ti-hɛ:ʔ/ 'to be torn' (S. Watson 1966: 22)
- (44) Sre (Bahnaric): /ha:l/ 'to cut' → /gə-ha:l/ 'to be cut' (Manley 1972: 46)

29.3.2.5 *Pronominal Prefixes*

In various Mon-Khmer languages, morphology involving pronouns shows paradigms, though these are almost entirely lexicalized and incomplete. For instance, in the Bahnaric language Koho, a homorganic nasal /N-/ derives interrogative pronouns, deictics, and other pronominals.

- (45) Koho (Bahnaric): /ci/ ‘thing’ → /ʔn-ci/ ‘what’; /kar/ ‘remaining’ → /ʔn-kar/ ‘how much remaining’; /te:ŋ/ ‘place’ → /ʔn-te:ŋ/ ‘where’; /vɛ/ ‘similar to’ → /ʔm-vɛ/ ‘like what’ (Hoàng Tuệ, et al. 1986: 53)

Such morphological evidence is even seen in Vietnamese, in which demonstratives have been shown to have at least a partial paradigm of initials, rhymes, and even tones, though these involve only initial consonants and not prefixes (Thompson 1965: 142).

As for deictics, distals are differentiated by vowel length in Pacoh (Katuic) and tone in Vietnamese.

- (46) Pacoh (Katuic): /ʔntih/ ‘that’ → /ʔnti:h/ ‘that (farther)’ (Nguyễn Văn Tài et al. 1984: 46)
 (47) Vietnamese (Vietic): /kiə¹/ ‘that’ → /kiə²/ ‘that (farther)’ (Nguyễn Đình Hoà 1966: 207)

Among eastern Mon-Khmer languages, similitive affixes are attached to demonstratives.

- (48) Bahnar (Bahnaric): /nu/ ‘this’ → /nə-nu/ ‘like this’ (Banker 1964a: 112)
 (49) Ruc (Vietic): /ni¹/ ‘this’ → /kə-ni¹/ ‘like this’ (Nguyễn Văn Lợi 1993: 88)

29.3.2.6 *Aspectual Prefixes and Infixes*

A variety of aspect-like meanings can be expressed variously through prefixes, infixes, and reduplicants (cf. Section 29.4.2). While the forms in northern and eastern Mon-Khmer appear more lexical, idiosyncratic, and thus derivational in nature, the Aslian forms are quite possibly inflectional, though they are included for now. The degree of productivity varies. In Katu, in contrast, the morphology occurs only on some verbs. In Katu, the completive is marked by a set prefix, while the imperfective is marked by a /Cv-/ prefix with a copied initial consonant. Morphological aspect is most common in Aslian and can be applied quite freely.

- (50) Bahnar (Bahnaric): /kʰan/ ‘to say’ → /jə-kʰan/ ‘to have said’ (Banker 1964a: 112)
 (51) Katu (Katuic): /te:k/ ‘to tear’ → /sar-te:k/ ‘already torn’; /bil/ ‘to lose’ → /babil/ ‘to keep losing’ (Costello 2001: 121–3)
 (52) Kentakbong (Aslian): /ciʔ/ ‘to eat’ → /ʔu-ciʔ/ ‘has eaten’ (Omar 1976: 956);
 Kentakbong (Aslian): /sapoh/ ‘sweeps’ → /sənapoh/ ‘is sweeping’ (Omar 1976: 968)
 (53) Khmu (Khmuic): /cɲək/ ‘one being nods once’ → /ɲək-kɲ-ɲək/ ‘one being keeps nodding downwards’ (Svantesson 1983: 57)
 (54) Temiar (Aslian): /gəl/ ‘sit (completed)’ → /gl-gəl/ ‘sit (incomplete)’ → /bər-gl-gəl/ ‘currently sitting’ (Benjamin 2011: 15)

29.3.2.7 *Verbalizing Prefixes*

Though varying in semantico-syntactic properties, a broadly defined class of verbalizing prefixes is seen in several sub-branches. Some derive general verbs from nouns.

- (55) Bru (Katuic): /raŋih/ ‘breath’ → /ta-ŋih/ ‘to breathe’ (Hoàng Văn Ma and Tạ Văn Thông 1998: 73)
- (56) Kri (Vietic): /pla:jh/ ‘an arm span’ → /p-a-la:jh/ ‘to measure by arm spans’ (Enfield and Diffloth 2009: 46)

Some have speech-related semantic properties. Koho (Bahnaric) has a prefix used to derive several verbs from nouns referring to the sounds of calls to attract or scare away animals.

- (57) Khmu (Khmuic): /pú:p-pú:p/ ‘(sighing sound)’ → /km-pú:p-pú:p/ ‘to sigh’ (Svantesson 1983: 53)
- (58) Koho (Bahnaric): /tɛ/ ‘the sound of calling/summoning pigs’ → /və-tɛ/ ‘to call/summon pigs’ (Hoàng Tuệ et. al. 1986: 53)
- (59) Mang (Mangic): /ʔaw³/ ‘speech’ → /ʔɔ¹-ʔaw³/ ‘to speak’ (Nguyễn Văn Lợi et al. 2008: 144)

29.3.2.8 *Desiderative/Inclined Prefixes*

The desiderative, expressing intention or inclination, is marked by prefixes in a few branches of Mon-Khmer over a wide geographic range. They are not cognates among the branches and thus are not reconstructable. Thus, these do have the potential to be clitics or clitics that have become prefixes, but that needs further testing.

- (60) Bru (Katuic): /taʔ/ ‘to do’ → /ʔi-taʔ/ ‘wants to do/about to do’ (Nguyễn Văn Tài et al. 1984: 36–9)
- (61) Kentakbong (Aslian): /ciʔ/ ‘to eat’ → /maʔ-ciʔ/ ‘wants to eat’ (Omar 1976: 955)
- (62) Khasi (Khasic): /jun/ ‘to urinate’ → /sam-jun/ ‘urge to urinate’ (Nagaraja 1985: 26)

29.3.3 Other Sub-categories

Several other morphological categories are seen in only two sub-branches, but it is worth noting as additionally there are multiple languages with these categories, and additional research would most likely reveal more instances. Table 29.7 shows an overview of the specialized semantic functions of the categories summarized in this sub-section.

29.3.3.1 *Negation Prefixes*

Negation prefixes occur on verbs in some languages. In general, these occur only lexicalized on specific words and have very limited productivity, and so this category is included in this section rather than Table 29.3.

- (63) Bahnar (Bahnaric): /wih/ ‘to go home’ → /kə-wih/ ‘not go home’ (Banker 1964a: 116)
 (64) Khmer (Khmeric): /miən/ ‘have’ → /k-miən/ ‘not have’ (Huffman 1970: 412)
 (65) Mon (Monic): /ket/ ‘take’ → /h-k-w-et/ ‘not take’ (Jenny 2003: 56)
 (66) Pacoh (Katuic): /ʔij/ ‘want’ → /ʔa-ʔij/ ‘hate’ (R. Watson 2011: 229)

29.3.3.2 *Existential/Locative Prefixes*

These prefixes (though see R. Watson (2011) for arguments that these are clitics) create existential predicates generally from personal pronouns, demonstrative pronouns, or interrogative pronouns.

- (67) Bru (Katuic): /naj/ ‘this’ → /ʔn-naj/ ‘as for/on this’ (Nguyễn Văn Tài et al. 1984: 39)
 (68) Pacoh (Katuic): /hɛ:/ ‘we’ → /ʔn-hɛ:/ ‘as for/on us’ (Nguyễn Văn Lợi, Đoàn Văn Phúc, and Phan Xuân Thành 1986: 42–3)
 (69) Ruc (Vietic): /ʔaj¹/ ‘who’ → /ʔa-ʔaj¹/ ‘as for/on one/whom’ (Nguyễn Văn Lợi 1993: 97)

29.3.3.3 *Involuntary Prefixes*

A phonologically similar /ta-/ prefix expressing non-intention to perform an action is seen in neighboring Katuic and Bahnaric languages.

- (70) Chrau (Bahnaric): /pəm/ ‘to hit’ → /ta-pəm/ ‘to accidentally hit’ (Thomas 1969: 98)
 (71) Katu (Katuic): /lə:j/ ‘throw’ → /ta-lə:j/ ‘throw accidentally’ (Costello 1966: 72)
 (72) Pacoh (Katuic): /ca/ ‘to eat’ → /ta-ca/ ‘to eat accidentally’ (S. Watson 1966: 23)

29.3.3.4 *Plural Stative Verbs*

Prefixes in Katuic and Khmuic mark adjectival verbs to indicate that the subjects are plural.

- (73) Pacoh (Katuic): /kɛt/ ‘small’ → /ka-kɛt/ ‘small (in regards to a plural subject)’ (Alves 2004: 39)
 (74) Khmu (Khmuic): /cɲák/ ‘one being nods once’ → /ĩ-ɲák/ ‘many beings nod once at the same time’ (Svantesson 1983: 57)

29.3.3.5 *Pretence*

In two geographically separate sub-branches, there are verbal affixes expressing the idea that an agent is pretending to perform an action.

- (75) Pacoh (Katuic): /taʔ/ ‘to work’ → /taʔ-n-taʔtaʔ/ ‘to pretend to work’ (S. Watson 1966: 29)
 (76) Palaung (Palaungic): /vy:n/ ‘to be drunk’ → /pən-vy:n/ ‘to pretend to be drunk’ (Milne 1921: 74)

Table 29.8 Number of days/years in Pacoh (Katuic)

Gloss	Form	Gloss	Form
yesterday	ʔiŋaj-ʔino:	last year	kumə:-ʔino:
2 days ago	ʔiŋaj-ʔntro:	2 years ago	kumə:-ʔntɾa:
3 days ago	ʔiŋaj-ʔntre:	3 years ago	kumə:-ʔntɾe:
4 days ago, etc.	ʔiŋaj-ʔnɾyən	4 years ago, etc.	kumə:-ʔnɾyən
tomorrow	ʔiŋaj-pəɾno:	next year	kumə:ŋ-ʔnnəh
2 days later	ʔiŋaj-pərra:	2 years later	kumə:-kumə:
3 days later	ʔiŋaj-pərrɛ:	3 years later	kumə:-kumɛ:
4 days later, etc.	ʔiŋaj-pərryən	4 years later, etc.	kumə:-kumyən

Source: R. Watson (1976).

29.3.3.6 Number of Days/Years

In Katuic and Palaungic, the amount of time before or after in terms of days or years can be expressed by additional morphology. In the sample from Pacoh (Katuic), the additional morphological material is suffix-like in a reduced phonological form of the historical source (Table 29.8).⁵ Reduced forms of numbers are clearly the source of this system (e.g. /pɛ:/ ‘three’ > /-ʔntɾɛ/, /pɹyən/ ‘four’ > /-ʔnɾyən/ etc.)

Palaung has a less complex system but does show a general /ʔa.kV-/ shape for number of days in the future.

- (77) Palaung (Palaungic): /ʔa.ky:/ ‘the day after tomorrow’; /ʔa.ke:/ ‘in three days’; /ʔa.kĩ:/ ‘in four days’; /ʔa.kũ:/ ‘in four days’ (Milne 1921: 95)

29.4 REDUPLICATIVE WORD FORMATION IN
MON-KHMER

Reduplication in Mon-Khmer most often involves (a) both copying and alternation of phonological material (e.g. rhyming, alliteration, and other types) and (b) the derivation of words which are vivid and semantically complex (e.g. an adjective referring

⁵ These may also be considered pseudo-compounds (cf. Section 29.5.1) as they appear to be closer phonologically to their lexical sources. That would also avoid the apparent conclusion that these involve suffixes, which is extremely typologically marked in Mon-Khmer. Another possibility is that the numbers are the heads, with the other material prefixes.

specifically to a tortuously winding road or the stumbling movement particular to someone who is drunk). Diffloth (1976a) has posited a special lexical class of ‘expressives’ insofar as these are syntactically distinct from other lexical classes as well as having various shared phonological and semantic properties. Onomatopoeia is one example of the kind of semantic function of reduplication in Mon-Khmer languages that functions in a special syntactic class, as noted by Svantesson (1983: 56). However, throughout the family, there is still a good deal of variety in the phonological patterns of reduplication in terms of its interaction with affixal morphology, the types of derived words, and its semantico-syntactic forms and features.

Due to the high degree of productivity of reduplication in Mon-Khmer varies and, based solely on grammatical descriptions, it is impossible to quantify this precisely. Thus, the following discussion is necessarily an impressionistic view. In many Mon-Khmer languages, there are numerous lexicalized forms created through alternating reduplication, with the creation of new terms being somewhat constrained but clearly available to allow for the hundreds or even thousands of forms in these languages. Moreover, there are instances of Mon-Khmer languages with open reduplicative processes. Little is known about the number of lexicalized reduplicants in individual Mon-Khmer languages, though a 1995 dictionary exclusively devoted to Vietnamese reduplicants contains over 5,000 entries (Viện Ngôn Ngữ Học (Institute of Linguistics) 1995). By comparison, in a general lexicon of Pacoh with about 10,000 entries, over 1,200 are reduplicants (Richard Watson p.c.). These significant percentages of lexicons are fairly representative of the significant productivity of this word-formation process. Moreover, both languages have other open reduplicative processes, even involving syntactic structures, as discussed in the following subsections.

29.4.1 Phonology

In this chapter, the term “full reduplication” means that the base and reduplicant are completely the same, while “alternating reduplication” means that the copying is partial, with some material copied and some material different from the base form. Alternating reduplication (variously referred to in studies of these languages as “partial reduplication,”⁶ “modified reduplication” (Diffloth 1979), “altering reduplication” (cf. Thomas 1971), “chameleon affixes” (cf. Thompson 1965), among other terms) can also be described in terms of what alternates. Table 29.9 shows several main phonological types of reduplication based on both what is copied and what alternates. Samples for each category are shown for reduplication of monosyllables in Vietnamese (Nguyễn Đình Hoà 1966, Viện Ngôn Ngữ Học 1995) and reduplication of bisyllabic words in other

⁶ The term ‘partial reduplication’ is used in descriptions of other languages in the world. However, as those often refer to copying of syllables of polysyllabic forms, that term is not considered precise enough.

Table 29.9 Types of alternating reduplication with monosyllabic bases

Types	Monosyllabic	Bisyllabic
Copying of initial / alternation of rhyme (i.e. alliteration)	<i>/xuk^{p3}/ (khúc) 'a stretch (of road)'</i> → <i>/xuk^{p3}-xwiw⁵/ (khúc khuỷu)</i> 'winding (of a road or stretch of land)'	Mang (Mangic): <i>/ʔuŋtip⁶/ 'cloud'</i> → <i>/ʔuŋtip⁶-ʔuŋtac⁶/ 'cloudiness'</i> (Nguyễn Văn Lợi et al. 2008: 162)
Copying of initial and final / alternation of main vowel	<i>/læk³/ (lắc) 'to move back and forth'</i> → <i>/luk^{p3}-læk³/ (lúc lắc) 'to swing/ move about continuously'</i>	Khasi (Khasic): <i>/jirwit/ 'crookedly' → /jirwit- jirwat/ 'indirectly' (Nagaraja 1985: 40)</i>
Copying of initial and main vowel / alternation of the final consonant	<i>/wan²/ (oằn) 'to be curved' → / wan²-waj⁶/ (oằn oại) 'to writhe in pain continuously'</i>	Koho (Bahnaric): <i>/sərvak/ 'good at, quick' → /sərvak-sərvan/ 'to be expert' (Tạ Văn Thông 1998: 199)</i>
Coda copying / alternation of initial (i.e. rhyming)	<i>/mɔ²/ (mò) 'to grope for' → / tɔ²-mɔ²/ (tò mò) 'to be curious about something'</i>	Chrau (Bahnaric): <i>/dichok-dipok/ 'calling a dog' (Thomas 1971: 157)</i>
Copying of base / alternation of tone	<i>/biəm³/ (bướm) 'butterfly' → / biəm¹biəm³/ (bướm bướm) 'butterfly'</i>	NA
Copying of segment into prefix	NA	Pacoh (Katuic): <i>/kia/ 'to saw' → / kaka/ 'to saw completely' (S. Watson 1966)</i>
Infix between base and reduplicant	NA	Koho (Bahnaric): <i>/k^het/ 'slow' → / k^het-rək^het/ 'slowly' (Tạ Văn Thông 1998: 208)</i>

Mon-Khmer languages. In those, the presyllables are copied completely, while the alternation occurs only in the stressed, second syllables.

Regarding the direction of reduplication, in alternating reduplication, the reduplicant tends to follow the root, but there are cases in which copied segments precede roots. By definition, prefixes with reduplicated material take that material from the following root syllable. Moreover, while base forms can be identified in many words derived from reduplications, there are numerous instances in which there is no identifiable base as neither syllable in what appears to be phonological reduplication has meaning on its own. This can be due to the loss of the root as a distinct meaning, losing its status as a free form, but it is also likely that reduplication itself has become a way to form new words even without a root source.

An analogical series of forms can occur without an actual (or at least uncertain) root source of the reduplication, as in Vietnamese in (78). These reduplicants share phonological and semantic properties of the forms (i.e. the syllable /luŋ^m/ (though with different tones) and the general meaning of instability), but there is no clear, unbound lexical source.

- (78) /luŋ^{m5}-laj¹/ (*lúng lay*) ‘to be shaking, unsteady’: /luŋ^{m1}-tuŋ^{m1}/ (*lung tung*) ‘to be in disorder, haphazardly done’: /luŋ^{m5}-tuŋ^{m5}/ (*lúng túng*) ‘confounded, awkward’: /luŋ^{m5}-t^huŋ^{m5}/ (*lúng thúng*) ‘to be long and loose (of clothes)’ (Nguyễn Đình Hòa 1966: 257)

Reduplication can interact with affixation in complex ways. Among the Aslian languages, there is a phenomenon known variously as a “incopyfixation” (Matisoff 2003), “coda copying” (Kruspe 2004), and “infixed reduplication” (Diffloth 1976a). In these cases, consonants from the coda are copied in infixes.

- (79) Temiar (Aslian): /mənə:j/ ‘opposite-sex sibling-in-law’ → /mɛ-j-nə:j/ ‘to be siblings-in-law’ (Benjamin 2011: 23)
 (80) Cheq Wong (Aslian): /hwac/ ‘whistle’ → /h-c-wac/ ‘whistling’ (Matisoff 2003: 31)

Reduplication can also involve additional complex polysyllabic reduplication, leading to even longer units of reduplicative phrases. The meanings are roughly the same, but there is a sense of emphasis.

- (81) Khasi (Khasic): /jram/ ‘suddenly’ → /jram-ram-ram/ ‘all of a sudden’; (Nagaraja 1985: 41)
 (82) Mnong (Bahnaric): /ʔuk-lak/ ‘disorderly’ > /luk-lak-tak-taj/ ‘disorderly’ (Dinh Le Thu 2007: 61)
 (83) Vietnamese (Vietic): /luŋ^{m5}-kuŋ^{m5}/ (*lúng củng*) ‘cumbersome/awkward’ → /luŋ^{m5}-ka¹-luŋ^{m5}-kuŋ^{m5}/ (*lúng ca lúng củng*) ‘cumbersome/awkward’ (Nguyễn Tài Căn 1996: 132)

Reduplication can also involve phrases up to the level of the sentence, a phenomenon best documented in Pacoh (Katuic) and Vietnamese (Vietic). In all cases, reduplicated material can be separated by repeated words. In (85), Vietnamese copies the initial onto a /-iək³/ rhyme.

- (84) Pacoh (Katuic): /papi/ ‘to speak’ → /papi-papoj/ ‘to talk and talk’
 pɛ ʔɪŋaj ɲa:j papi ɲa:j papoj
 three day they speak they speak-RED
 ‘Three days they talk and talk.’ (R. Watson 1966a: 180)
 (85) Vietnamese (Vietic): /ɲu⁵/ (*ngủ*) ‘to sleep’ → /ɲu⁵-ɲiək³/ (*ngủ ngủiéc*) ‘to sleep (dismissive)’
 Tại sao con cứ khóc hoài vậy không ngủ không ngủiéc bao giờ
 why child just cry incessantly thus NEG sleep NEG C+iək ever
 ‘Why are you crying all the time? You’re never sleeping or anything!’
 (Vu 2007: 194)

29.4.2 Semantic and Syntactic Functions of Reduplication in Mon-Khmer

Reduplication in Mon-Khmer is primarily non-grammatical, deriving content vocabulary but with certain semantic properties. However, in a number of instances, there are grammar-like functions, primarily adverbial functions expressing degree of intensity, plurality, or time. Reduplication in Vietnamese has been shown to express distribution, iteration, attenuation (reduced force), intensification, and emphasis (Thompson 1965: 151–60). Diffloth (1976a) has used the term “expressives” to refer to special syntactic distributional properties of many of these reduplicants. For this study, they will be grouped with core parts of speech (nouns, verbs, adjectives, and adverbs) as much as they overlap with the function and distribution of those, with the question of the specialized distribution of “expressives” left aside. It is worth noting that while Vietnamese has no affixes, its reduplication has a notable number of morpho-semantic functions, as shown in Table 29.10.

29.4.2.1 Time Adverbs

A range of general adverbs is derived, often resulting in adverbials of extended periods of time.

- (86) Mang (Mangic): /lin¹cɔ³/ ‘long’ → /lin¹cɔ³-lin¹cɔ³/ ‘continuously’ (Nguyễn Văn Lợi et al. 2008: 160)

Table 29.10 Specialized semantico-syntactic functions of reduplication in Mon-Khmer

Sub-branch categories	KS	MG	PL	KU	AS	MN	KR	BN	KT	VT	VN
1. Time adverbs	+	-	+	+	-	+	-	+	-	+	+
2. Degree of intensity	-	+	+	+	+	-	-	+	-	+	+
3. Iteration/progression	-	+	-	-	+	-	-	+	+	+	-
4. Generalized	-	+	-	-	-	-	-	+	+	-	+
5. Derogatory/dismissive	-	+	-	-	-	-	-	+	+	-	+
6. Distributive/plural nouns	-	-	-	-	-	-	-	+	-	+	+
Total number	1	4	2	2	2	1	0	6	3	4	5

Note: the following abbreviations are used for the various sub-branches: AS-Aslian, BN-Bahnaric, KR-Khmeric, KS-Khasic, KT-Katuic, KU-Khmuic, MG-Mangic, MN-Monic, NC-Nicobarese, PL-Palaungic, VN-Vietnamese, VT-Vietic.

- (87) Vietnamese (Vietic): /luənɪ / (*luôn*) ‘often’ → /luənɪ-luənɪ/ (*luôn luôn*) ‘very often/always’ (Nguyễn Đình Hoà 1966: 258)
 (88) Palaung (Palaungic): /shɛ:-shɛ:/ ‘always’ (Milne 1921: 93)
 (89) Sre (Bahnaric): /rəhɑ:j/ ‘gradually’ → /rəhɑ:j-rəhɑ:j/ ‘gradually’ (Manley 1972: 40)

29.4.2.2 *Degree of Intensity*

Expressing degree of intensity through manner adverbs is a common function of reduplication, often done through full reduplication. There is in some languages a lightening or slight reduction in the intensity.

- (90) Koho (Bahnaric): /jɔʔ/ ‘a long time’ → /jɔʔ-jɔʔ/ ‘a little long’ (Tạ Văn Thông 1998: 186)
 (91) Mnong (Bahnaric): /daŋ/ ‘hard’ → /da-daŋ/ ‘a little hard’ (Dinh Le Thu 2007: 63)
 (92) Muong (Vietic): /haw³/ ‘to want’ → /haw³-haw³/ ‘to want somewhat’ (Hoàng Văn Hành 1998: 168)

Reduplication in Mon-Khmer can conversely express stronger intensification. Note that this is accomplished through a variety of reduplicative types in examples. Mnong (Bahnaric) uses a prefix in (91) above but an infix plus reduced reduplicant in (95) below.

- (93) Khasi (Khasic): /kʰruj/ ‘in many’ → /kʰruj-ruj-ruj/ ‘in many great numbers’ (Nagaraja 1985: 41)
 (94) Mang (Mangic): /dɔ¹/ ‘much’ → /dɔ¹-dɔ¹/ ‘very much’ (Nguyễn Văn Lợi et al. 2008: 160)
 (95) Mnong (Bahnaric): /preh/ ‘tall’ → /preh-ʔa-reh/ ‘very tall’ (Dinh Le Thu 2007: 63)
 (96) Vietnamese (Vietic): /ɲiw²/ ‘a large amount’ → /ɲiw²-ɲiw¹/ ‘a rather large amount’ (Vu 2007: 171)

29.4.2.3 *Generalized Meaning*

Reduplication of verbs can have a generalizing meaning, referring to the action in a general way rather than a specific instance of the action. In such cases, the verbs are less likely to be transitive, and the reduplicated nouns are not specific items, but rather the class of items and thus less likely to be definite.

- (97) Bru (Katuic): /rah/ ‘to hunt (something)’ → /rurah/ ‘to hunt (in general)’ (Hoàng Văn Ma and Tạ Văn Thông 1998: 77)
 (98) Koho (Baharic): /ʔmpʰa/ ‘pants’ → /ʔmpʰa-ʔmpʰuŋ/ ‘pants (in general)’ (Tạ Văn Thông 1998: 199)
 (99) Mang (Mangic): /əa¹/ ‘to eat’ → /əa¹əac⁶/ ‘to eat (in general)’ (Nguyễn Văn Lợi et al. 2008: 164)

29.4.2.4 *Derogatory/Dismissive Meaning*

A derogatory or dismissive meaning is expressed by alternating reduplication, with a geographic clustering in the east in Vietnam. Forms to express this semantic category

tend to involve open production processes, as is the case in Vietnamese with its /C₁-iək⁵/ post-root reduplicant, as in (104). This category of reduplication in these languages may involve a bundling of additional semantic features and functions. For example, in Katu (Katuic) in (102), the initial of the root is copied with a /-ə/ replacing the rhyme to indicate a dismissive negation of verbs or nouns.

- (100) Bahnar (Bahnaric): /kələŋ/ ‘to look at’ → /kələŋ-kələj/ ‘not worth looking at but look anyway’ (Banker 1964b: 121)
- (101) Halang (Bahnaric): /kuuj/ ‘to sleep’ → /kuuj-kajʔ/ ‘to sleep on the ground with nowhere else to sleep’ (Cooper 1966: 32)
- (102) Katu (Katuic): /do:ŋ/ ‘house’ → /do:ŋ-də/ ‘no house whatsoever’; /kəra:ŋ/ ‘to worry’ → /kəra:ŋ-kərə/ ‘not worrying whatsoever’ (Nguyễn Hữu Hoàng 1998: 152–53)
- (103) Mang (Mangic): /to⁴/ ‘big’ → ‘big (dismissive)’ (Nguyễn Văn Lợi et al. 2008: 164)
- (104) Vietnamese (Vietic): /bæn⁶/ (*bạn*) ‘friend’ → /bæn⁶-biək³/ (*bạn biếc*) ‘friends (dismissive)’ (Vu 2007: 174)

29.4.2.5 *Distributive/Plural Nouns*

In a type of iconic word formation, full reduplication of nouns and measure words often indicates plurality or a kind of distributive meaning. This is a rather universal function of full reduplication, also seen in Chinese and Malay. It is common for quantifiers with distributive meanings (e.g. “each,” “all,” etc.) to accompany these nouns.

- (105) Koho (Bahnaric): /naʔ/ ‘person’ → /naʔ-naʔ/ ‘some people’ (Tạ Văn Thông 1998: 196)
- (106) Mnong (Bahnaric): /ʔntok / ‘place’ → /ʔntok-ʔntok/ ‘every place’ (Dinh Le Thu 2007: 62)
- (107) Vietnamese (Vietic): /ɲæ²/ (*nhà*) ‘house’ → /ɲæ²-ɲæ²/ ‘all the houses’ (Thompson 1965: 151)

29.5 COMPOUNDING

Mon-Khmer languages are strongly right-branching syntactically (with some specific, restricted exceptions), and as expected, lexical compounding is as well, with left-headed lexical compounds in most cases (Table 29.11). Compounding among Mon-Khmer languages most often follows the head-dependent pattern, though dual-head compounds are also a common word-formation pattern. Compounding does appear more often north of the Malay peninsula as Aslian languages appear to rely less on this word-formation strategy (Matisoff 2003: 55), while conversely, compounds are extremely common

Table 29.11 Types of lexical compounds in Mon-Khmer

Compound type	Examples
Noun plus noun	Katu (Katuic): /dak-mət/ (water-eye) 'tears' (Nguyễn Hữu Hoàn and Nguyễn Văn Lợi 1998: 143) Khasi (Khasic): /ʔum-mət/ (water-eye) (Nagaraja 1985: 59) Khmu (Khmuic): /ʔòm-mət/ (water-eye) (Premrsirat 1987: 29) Muong (Vietic): /dak ³ -mat ¹ / (water-face) 'tears' (Nguyễn Văn Khang et al. 2002: 142) Nicobarese (Nicobarese): /dak-mət/ (water-eye) (Man 1889: 95)
Noun plus verb/ adjective	Khmu (Khmuic): /ʔòm-kséh/ (water-to fall) 'waterfall' (Premrsirat 1987: 29) Khasi (Khasic): /don-kam/ (is/exist-work) 'to need' (Nagaraja 1985: 66) Koho (Bahnaric): /sa-təɽəm/ (eat-dishonest) 'to cheat' (Hoàng Tuệ et. al. 1986: 43) Pacoh (Katuic): /cɛ:t-tarma/ (die-sick) 'to die of sickness' (Nguyễn Văn Lợi et al. 1986: 135)
Verb plus verb/ adjective	Khmu (Khmuic): /sɨh-hó:j/ (lie-play) 'lie down for a rest' (Premrsirat 1987: 29) Katu (Katuic): /dak-mət/ (water-eye) 'tears' (Nguyễn and Nguyễn 1998: 143) Ruc (Vietic): /khik ³ -kitəŋ ³ / (healthy-young) 'robust' (Nguyễn Văn Lợi 1993: 73) Vietnamese (Vietic): /nɔj ³ -cəj ¹ / (nói chơi) (to speak-to play) 'to joke' (Nguyễn Đình Hòa 1966: 292)
Verb plus noun	Bru (Katuic): /taʔ-kol/ (do-mouth) 'to speak' (Hoàng Văn Ma and Tạ Văn Thông 1998: 91) Khasi (Khasic): /ʃaʔ-ʃkor/ (allow-ear) 'to listen' (Nagaraja 1985: 66) Khmu (Khmuic): /wá:r-lù:l/ (hot-stomach) 'hungry' (Premrsirat 1987: 29) Li (Mangic): /bi33-mən55/ (buy-wife) 'to marry (a woman)' (Li 1996: 143)

(thousands of words) in the Vietnamese lexicon, which utilizes extensively a two-syllable compounding strategy similar to that of Chinese.

29.5.1 Compounds with Dependents

Right-branching phrase structure (e.g. head-modifier, verb-complement, preposition-noun, etc.) is common to all Mon-Khmer languages, and such is also the case in compounding. Compounding can be categorized according to the combinations of parts of speech, including noun-noun, noun-verb/adjective, verb-verb, and verb-noun compounds. There are semantically transparent, predictable combinations and those with meanings that cannot be predicted by the individual elements. Examples (103) to (107) are included to show an instance of a geographic calquing also involving cognates.

One concept regarding the nature of compounding in Southeast Asia is that of 'pseudo-compounding' (cf. Thompson 1965, Starosta 1992), in which the morphemes may have identifiable, distinct meanings but which are bound forms or which have broader semantic functions, for instance, a word meaning "house" referring to any buildings, as in the compounds in (108) in Katu (Katuic).

- (108) Katu (Katuic): /doŋ-rahɔk/ (house-study) ‘school’; /doŋ-pajəh/ (house-to cure) ‘hospital’; /doŋ-dak/ (house-water) ‘housing/household’ (Nguyễn Hữu Hoàn and Nguyễn Văn Lợi 1998: 130)

Phonological specialization of pseudo-compounds is most evident in Khasi (Khasic). For instance, in Khasi, “tear” is /ʔum-mat/ (‘water-eye’), in which the compound contains a shorter form for “eye” than the unbound form /kʰmat/ ‘eye’ (Nagaraja 1985: 59). Similarly, in Khasi the three-member compound /sop-ti-kti/ (cover-hand-hand) ‘blouse’ contains /ti/, the reduced form of /kti/ ‘hand’ (Nagaraja 1985: 68). In other instances, compounds may have one identifiable morph with another form which is bound and semantically non-distinct, such as /durkʰmat/ ‘face’ consisting of the bound /dur-/ and /kʰmat/ ‘eye’ (Nagaraja 1985: 71). The generalized semantics and phonological changes of the bound forms is suggestive of something other than full compounds.

Pseudo-compounds are also seen in the Sino-Vietnamese lexicon, which are also notable as they are right-headed. Right-headed pseudo-compounds are seen in the thousands of Chinese loanwords in Vietnamese as well as Muong, though in perhaps smaller numbers. The two syllable forms may both have distinct meanings, though they are bound morphemes and cannot function alone as words.

- (109) Mandarin Chinese: /gɔŋ¹-bu⁴/ (*gōng bù* 公布) (public-announce) ‘to make public/announce’
 (110) Vietnamese (Vietic): /kawŋ^{m1}-bo³/ (*công bố*) (public-announce) ‘to make public/announce’ (Nguyễn Đình Hoà 1966)
 (111) Muong (Vietic): /coŋ¹-po⁵/ (public-announce) ‘to make public/announce’ (Nguyễn Văn Khang, Bùi Chỉ, and Hoàng Văn Hành 2002: 410)

Grammatical compounds common to many Austroasiatic languages include pronouns and numbers. Compounding in the formation of pronouns is widespread and manifested in various ways (cf. Pinnow 1965).

- (112) Bugar (Mangic): /wi³¹-biɔ³¹/ (we (inclusive)-two people) ‘we two’ (Li 1996: 147–8)
 (113) Pacoh (Katuic): /ʔndɔ-ʔapɛ/ (possessive-3p) ‘theirs’ (S. Watson 1964: 85)
 (114) Vietnamese (Vietic): /kæk³-ʔawŋ^{m1}/ (plural-gentleman) (*các ông*) ‘you (plural) gentlemen’ (Thompson 1965: 183)

29.5.2 Dual-headed Compounds

Dual-headed compounds appear to be more concentrated in the eastern regions where Mon-Khmer languages are spoken, with a few exceptions. This kind of compounding most often involves nouns, but active verbs and even stative verbs are also seen. Some common types of dual-headed compounds indicate family members (e.g.

“mother-father” to mean “parents,” “wife-husband” to mean “married couple,” “younger sibling-older sibling” to mean “siblings in general,” etc.) as well as general categories of items (e.g. “pants-shirt” to mean “clothing,” “table-chair” to mean “furniture,” etc.).

- (115) Chrau (Bahnaric): /jun-jikɛ/ (deer-wild pig) ‘destructive animals’ (Thomas 1971: 141)
 (116) Mang (Mangic): /θa⁵-θɔ¹/ (cooked rice-taro) ‘food’ (Nguyễn Văn Lợi et al. 2008: 132)
 (117) Vietnamese (Vietic): /bæn²-ɣe³/ (table-chair) (*bàn ghế*) ‘furniture’ (Nguyễn Đình Hoà 1966: 9)

In some cases, there are two words with the same meaning referring to the general category.

- (118) Koho (Bahnaric): /ha:ŋ-vənɛ/ (pretty-pretty) ‘pretty’ (Hoàng Tuệ et. al. 1986: 41)
 (119) Katu (Katuic): /ʃaj-pɔr/ (help-help) ‘to help’ (Nguyễn Hữu Hoàn and Nguyễn Văn Lợi 1998: 135)
 (120) Vietnamese (Vietic): /sɿŋ¹-dɛp⁶/ (*xinh đẹp*) (pretty-pretty) ‘pretty’ (Nguyễn Đình Hoà 1966: 558)

Generalized compounds have been shown to be syntactically separable, with reduplicated syntactic material in Mal (Khmuic) and Ngeq and Pacoh (Katuic), somewhat like separable reduplicants in Vietnamese and Pacoh (Section 29.4.1).

- (121) Mal (Khmuic): /lɔʔ-lɛ:m/ (good-good) ‘good’
 nam kaj^h p^hjam lɔʔ p^hjam lɛ:m
 he is person good person good
 ‘He is a perfect person.’ (Filbeck 1996: 92)
 (122) Ngeq (Katuic): /krɛ-cuəŋ/ (vine-vine) ‘vines’
 ɲa:j jɔ:ʔ su krɛ su cuəŋ
 they go search vine search vine
 ‘They days they talk and talk.’ (Smith 1973: 103)

29.6 CONCLUDING THOUGHTS

The data presented in this chapter suggest a wider array of morphological features than previously considered common in Mon-Khmer. While Mon-Khmer morphology is largely fossilized, derivational morphology, there is a good deal of productive morphology. Semantic functions such as causation and reciprocity are created through affixation, but also inclination and state. Reduplication creates iconic meanings of distribution and

repletion, but it is also used to express displeasure and general categories of things. In terms of geography, it must be noted that the highest numbers of morphological function are on the geographic edges of Mon-Khmer (e.g. Aslian, Khmuic, and Katuic), with smaller numbers in the central region (e.g. Mon and Khmer). Beyond the categories identified in this chapter, there are dozens of other morphological phenomena that are in individual languages or sub-branches not included in this study but which further the notion that the morphology is richer than supposed.

CHAPTER 30

AUSTRONESIAN

ROBERT BLUST

30.1 THE AUSTRONESIAN LANGUAGE FAMILY

ALTHOUGH Indo-European and Sino-Tibetan have more speakers, the Austronesian (AN) language family ranks second in number of languages, with an estimated 1,257 members spread over 206 degrees of longitude from Madagascar to Easter Island, and 72 degrees of latitude from northern Taiwan to southern New Zealand (Lewis 2009). It is now generally agreed that the greatest genetic diversity within AN is found among the fifteen remaining aboriginal languages of Taiwan, which fall into at least nine primary branches of the family, while all AN languages outside Taiwan belong to a single enormous sub-group called ‘Malayo-Polynesian’ (Blust 2009) (Figure 30.1).

The largest, and therefore best-studied, AN languages are found in western Indonesia and the Philippines. These include Bahasa Indonesia (based on Malay), the national language of the Republic of Indonesia, which is spoken by well over

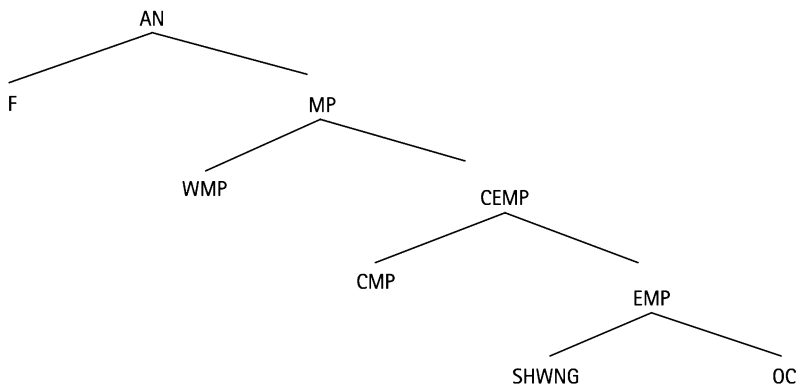


FIGURE 30.1 The Austronesian family tree

90% of a population that numbered just under 242 million in July 2005, and Filipino (based on Tagalog), the national language of the Republic of the Philippines, also spoken by over 90% of the population, which was not quite 88 million in July 2005 (CIA-The World Factbook). In addition, several languages of Indonesia and the Philippines that lack official status are exceptionally large by general AN standards, most notably Javanese, with an estimated 85 million first language speakers, and Sundanese (west Java), with an estimated 34 million first language speakers in 2000. In sharp contrast, no AN language of eastern Indonesia or the Pacific has as many as a million speakers, and many language communities, particularly in Melanesia, range from 100 to 300 speakers, numbers that apparently have been sustained at these levels for many centuries.

30.2 WORD-FORMATION PROCESSES IN AUSTRONESIAN LANGUAGES

Given the size and geographical spread of AN, as well as noteworthy contact influences from Austroasiatic and Tai-Kadai languages on the Asian mainland, and Papuan languages in the region of New Guinea, it might be expected that languages in this family exhibit a wide range of typological variation, making it difficult to generalize. Moreover, since the theoretical focus of this volume is on derivational processes it must be made clear from the outset that the inflection/derivation distinction is far less clear in many AN languages than it is in languages such as English. Keeping these caveats in mind, I will attempt here to broadly outline some of the major types of word-formation processes found in AN languages.

30.2.1 Voice/“Focus” Morphology

Proto-Austronesian (PAN) had a typologically unusual verb system in which an actor voice (or “actor focus”) marked by *-um- was matched by three “passive” voices (not everyone uses this terminology), a “direct passive” marked by *-ən, a “local passive” marked by *-an, and an “instrumental/benefactive passive” marked by *Si- (Wolff 1973). The perfective infix *-in- (prefixed to vowel-initial bases) could co-occur with all of these affixes except *-ən, which had a zero allomorph in perfective verbs (*k<um>aən ‘to eat,’ *k<um><in>aən ‘ate,’ but *kaən-ən ‘be eaten by someone; food,’ k<in>aən ‘was eaten by someone,’ etc.). The main outlines of this system are preserved in many so-called “Philippine-type” languages, including most languages of Taiwan, the Philippines, northern Borneo, and northern Sulawesi, as well as Chamorro of western Micronesia, and Malagasy. While some analyses call these affixes inflectional, others see them as derivational, and the words marked by *-ən, *-an, and *Si- frequently function as both

nouns and verbs, as seen in (1a–b) from Thao of central Taiwan (PV.PERF = patient voice perfective; SA = un glossable particle, LIG = ligature; Blust 2003):¹

- (1) a. *wazish in-apa sa suma*
 wild pig PV.PERF-carry SA someone
 ‘Someone carried the pig’
 b. *inay nak a in-apa wa aniamin*
 this/these my LIG PV.PERF-carry LIG thing
 ‘These are the things that I carried’ (lit. ‘these are my carried things’)

In (1a) *in-apa* seems clearly to be verbal, since otherwise the sentence would be verbless. However, in (1b) *in-apa* seems just as clearly to be nominal, since only nouns can be possessed. Each of these words adds *-in-* ‘perfective’ to the base *apa* ‘carry’, but the interpretation of the morphologically complex result as verb or noun is entirely dependent on the larger syntactic context in which the word is embedded. Examples such as this could be multiplied many times over in scores of Philippine-type languages.

30.2.2 Other Types of Segmental Derivational Affixes

Examples (1a–b) illustrate the difficulty of marking the noun/verb distinction purely by morphology. However, Philippine-type languages have rich morphological systems, and many affixes are clearly derivational. A few prominent examples should suffice to illustrate the general pattern. PAN **ma-* was added to nominal bases to create stative verbs, a process that persists in many Philippine-type languages, but is fossilized in a number of languages that have radically altered their grammatical type. An example is seen in Tagalog *sakit* ‘disease, sickness’: *ma-sakit* ‘sore, painful’, where the stative verb is derived from a nominal base by the prefix *ma-*, but the Tongan cognate *mahaki* ‘sick person, patient’ has lost the morpheme boundary and reinterpreted the reanalyzed base as a noun, which now must be verbalized by suffixation with *-ʔia*, as in *mahaki-ʔia* ‘to be sick or unwell’. Examples such as this show that over time derivational processes that were once highly productive can fall into desuetude, and yet survive as reanalyzed parts of longer words, which are then no longer obligated to conform to the word-class originally signaled by the now fossilized affix.

Austronesian languages make use of prefixes, infixes, suffixes, and circumfixes as instruments of word formation, and each of these affix types can be used to derive one word-class from another. An example of derivation via prefixation was just given.

¹ In Austronesian linguistics the term “ligature” or “linker” describes a particle, usually having the shape V or CV, which is required between a head and its attribute, as in Thao *ma-tubu wa saran* (stat-wet LIG road) ‘a wet road’, *nak a lharina* (1sg.poss LIG ear) ‘my ear’, *kawi a filhaq* (tree LIG leaf) ‘leaves of a tree’, or *tata wa shaba* (one LIG hundred) ‘one hundred’. In some languages, such as Tagalog, allomorphs of the ligature vary between CV and -C, as with *ma-baʔit na bátaʔ* (STAT-good LIG child) ‘a good child’, or *ʔúso-ʔ báboy* (snout-LIG-pig) ‘protruding upper lip’ (lit. ‘pig’s lip’).

Derivation via infixation can be seen in many of the languages of the western Solomon Islands in the Oceanic sub-group of AN, where the infix *-in-*, which probably reflects PAN **-in-*, but with loss of all reference to perfectivity, is a highly productive source of deverbal nouns. In Hoava, spoken on the island of New Georgia in the Solomons chain, Davis (2003: 39) says that “This affix is extremely productive and can be used with virtually any active or stative verb to create a noun.” This is shown by examples such as (i) **result of an action**: *asa* ‘grate’: *in-asa* ‘pudding made from grated cassava,’ *guzala* ‘twist bark to make string’: *g-in-uzala* ‘string made from twisted bark,’ *bukulu* ‘defecate’: *b-in-ukulu* ‘feces’; (ii) **object which undergoes an action**: *babana* ‘to tow’: *b-in-abana* ‘towed object,’ *gerigeri* ‘gather building materials’: *g-in-erigeri* ‘logs, sticks, vines, etc. needed for building a house,’ *mae* ‘come’: *m-in-ae* ‘people who have arrived’; (iii) **an abstract noun describing an action**: *dumi* ‘punch’: *d-in-umi* ‘act of punching,’ *hade* ‘wrap’: *h-in-ade* ‘act of wrapping,’ *boru* ‘to massage’: *b-in-oru* ‘massaging, cure’; (iv) **abstract nouns from experiential and stative verbs**: *edo* ‘happy’: *in-edo* ‘happiness,’ *hiva* ‘want’: *h-in-iva* ‘wishes, desires,’ *to* ‘alive’: *t-in-o* ‘life.’ In a few cases nouns may be created from existing nouns, and *-in-* may be inserted within a prefix to nominalize a derived verb, as with *va-bobe* ‘to fill’ (with causative *va-*): *v-in-a-bobe* ‘filled object,’ *vari-razae* ‘fight each other (with reciprocal *vari-*): *v-in-ari-razae* ‘war, battle,’ *ta-poni* ‘be given’ (with passive *ta-*): *t-in-a-poni* ‘gift.’

Derivation via suffixation is found in a number of the languages of western Indonesia, as Malay/Bahasa Indonesia, where the suffix *-an* is used to derive many nouns from other parts of speech, or from other nouns, as with *makan* ‘eat’: *makan-an* ‘food,’ *minum* ‘drink’: *minum-an* ‘a drink,’ *satu* ‘one’: *satu-an* ‘a unit,’ *puluh* ‘ten’: *puluh-an* ‘a multiple of ten,’ or *hari* ‘day’: *hari-an* ‘daily newspaper’ (Macdonald and Soenjono Darjowidjojo 1967). Similarly, in Kelabit of northern Sarawak, which has lost all verbal uses of the historical locative voice suffix **-an*, a reflex persists only to derive nouns of location, as with *irup* ‘what is drunk; way of drinking’: *rup-an* ‘drinking hole (as for animals in the jungle),’ *guta* ‘ford a river’: *guta-an* ‘fording place,’ or *dalan* ‘path’: *dalan-an* ‘path made by repeated walking over the same course’ (emphasis on location; Blust 1993).

A prominent type of derivation using the circumfix **ka-BASE-an* to mark an adversative passive is reflected in a number of languages, as in Pazeh, a recently extinct language of north-central Taiwan: *akux* ‘heat’: *ka-akux-an* ‘get heatstroke,’ *bari* ‘wind’: *ma-bari* ‘to blow, of the wind’: *ka-bari-an* ‘be blown about by the wind,’ *ma-kələm* ‘salty’: *ka-kələm-an* ‘oversalted, of cooked food,’ *udan* ‘rain’: *ka-udan-an* ‘be caught in the rain.’ A very similar derivational process is seen in Malay, Javanese, and some other languages of western Indonesia, as with Malay *hujan* ‘rain’: *kə-hujan-an* ‘be caught in the rain,’ *malam* ‘night’: *kə-malam-an* ‘benighted, overcome by darkness,’ *tidur* ‘sleep’: *kə-tidur-an* ‘fall asleep involuntarily (when intending to stay awake),’ or *sian* ‘bright daylight’: *kə-sian-an* ‘oversleep (wake up to bright daylight).’ In each of these cases words formed with a reflex of **ka-* *-an* describe an unfortunate or undesirable situation that results from agentless circumstances beyond the experiencer’s control (hence “adversative”).

In Philippine-type languages with their rich affix potential one and the same base can be transformed into a variety of related words through independent prefixation, infixation, suffixation, etc., or through simultaneous combinations of affixes. An example of

the first type is seen in Thao *tutu* ‘breast’ *pa-tutu* ‘to nurse a baby’: *t<m>utu/tutu-an* ‘to nurse, of a baby,’ where *pa-* creates a denominal transitive verb, and *-m-* and *-an* create intransitive counterparts *t<m>utu*, focused on the action, and *tutu-an* on the location, as shown in sentences (2a–c); AV = actor voice, LV = locative voice:²

- (2) a. *ina wa cicu pa-tutu sa azazak*
 mother LIG 3sg CAUS-breast SA child
 ‘His mother nursed the child’
 b. *azazak t<m>utu*
 child breast-AV
 ‘The child is nursing’
 c. *tutu-an sa azazak*
 breast-LV SA child
 ‘The child is nursing (at the breast)’

Space does not allow me to enter into detail (for more cf. Blust 2009, especially ch. 6, ‘Morphology’), but bases that host three or four affixes in one word are not unusual, particularly in Philippine-type languages, and in South Sulawesi languages such as Buginese or Makasarese, which have elaborated a rich innovative morphological system from the inherited Philippine-type voice (or “focus”) system (van den Berg 1996). Examples of bases hosting three affixes to form a single word are:

- (3) Ilokano (northern Philippines; Rubino 2000): *túrog* ‘sleep’: *m<in>aka-pa-túrog* ‘s.t. that caused s.o. to sleep’ (*maka-* ‘intransitive potentive,’ *-in-* ‘perfective,’ *pa-* ‘causative’)
 (4) Kadazan (north Borneo; Kadazan Dusun Cultural Association 1995): *patay* ‘die’: *mod-tig-ka-patay* ‘at the point of death’ (*mod-* ‘affected by,’ *tig* ‘immediate,’ *ka/ko-* ‘nominalizer’)
 (5) Makasarese (south Sulawesi; Cense 1979): *tawa* ‘share’: *at-tawa-əŋ-i* ‘half a share of this and half a share of that’ (*at* (?), *əŋ* ‘passive nominalizer,’ *-i* ‘3p?’)

Examples of bases hosting four affixes to form a single word are:

- (6) Ilokano: *liwá* ‘amusement; recreation’: *pag-pal-pa-liwa-án* ‘comfort, solace’ (*pag-* ‘instrumental,’ *pal* (?), *pa-* ‘causative,’ *-an* ‘nominalizer’)
 (7) Makasarese: *mate* ‘die, dead’: *an-tu-mate-aŋ-i* ‘to take responsibility for someone’s burial’ (*an-* ‘active verb,’ *tu-* ‘someone who,’ (*paŋ*)*X-aŋ* ‘place where one X,’ *-i* ‘agent’)

² Because Philippine-type languages allow multiple prefixes in the same word, as well as infixes, angled brackets are conventionally used to distinguish infixes from prefixes that are not initial. If a form like Ilokano *m<in>aka-pa-túrog* ‘something that caused someone to sleep’ were written *m-in-aka-pa-túrog*, for example, the morphological parsing of the form would clearly be more difficult, since an alternative reading of the latter representation would be a base with four prefixes.

Several observations can be made about this data. First, as can be seen, some affixes are difficult to gloss, even with relatively detailed published accounts of the languages in question. Second, in most languages the “piling up” of affixes occurs only, or primarily with prefixes. Third, given the general difficulty of distinguishing derivation from inflection in many AN languages there generally is no clear pattern in which derivational affixes are closer to the base while inflectional affixes are more peripheral. Fourth, unlike circumfixes, in which the meaning of the whole cannot be attributed to the meanings of the parts, the accumulation of multiple affixes follows an incremental pattern, thus suggesting cyclical addition of semantic components to the meaning of the whole. Finally, although polymorphemic words in most AN languages clearly remain words, in South Sulawesi languages such as Makasarese the line between word and sentence can become blurred, producing some single word propositions that are reminiscent of polysynthetic languages.

30.2.3 Subtractive Morphology

Subtractive morphology appears to be rare in the world’s languages, and has thus received scant attention in the generalist literature, but this pattern is found in a number of the languages of the Philippines and western Indonesia.

Stevens (1997) has drawn attention to what he calls ‘truncation phenomena’ in Bahasa Indonesia, by which he means the shortening of a base as a morphological process. He illustrates this in several domains, of which only kin terms and personal names need concern us here. In a number of AN languages the dropping of an initial segment or syllable marks the vocative form (Blust 1979), as shown in Table 30.1

Moreover, in some of the languages of Borneo the vocative of personal names is formed in a similar way, as in Timugon Murut of Sabah, where subtraction of an initial consonant or the entire first syllable can be used productively, as with *Ohn!* vocative of ‘John’ (D. J. Prentice, p.c.).

Table 30.1 Derivation by subtraction in the vocative forms of kinship terms

Language	Reference/address	Vocative	Gloss
Bikol	apoʔ	poʔ	grandparent, elder
Cebuano	nánay	nay	mother
Cebuano	tátay	tay	father, uncle
Malay	adek	dek	younger sibling
Malay	kakak	kak	elder sibling
Banjarese	adinj	dinj	younger sibling

30.2.4 Suprasegmental Affixes

Quite apart from affixation or subtractive morphology, both of which involve phonological segments as elements of word formation, a number of AN languages use stress shift to mark morphological distinctions. This is fairly common in languages of the northern and central Philippines which have phonemic lexical stress, but also use stress shift to mark categorial distinctions in word-class, as with Tagalog *básag* ‘a crack in earthenware, etc.’: *baság* ‘cracked; broken,’ *gísín* ‘waking up from sleep (n.):’ *gisín* ‘awake, not asleep,’ or *túlís* ‘point; sharp end’: *tulís* (given as the equivalent of *ma-túlís*) ‘pointed, sharp.’ A similar use of stress shift to derive adjectives or stative verbs, and occasionally other parts of speech, is also well-developed in Toba Batak of northern Sumatra, which generally has penultimate stress: *árga* ‘Bargain!’ (imper.): *argá* ‘expensive,’ *hójot* ‘Go quickly!’ (imper.): *hojót* ‘quick,’ *gógo* ‘Push hard!’ (imper.): *gogó* ‘strong,’ *ásom* ‘lemon’: *asóm* ‘sour,’ *díla* ‘tongue’: *dilá* ‘a big talker; person who exaggerates.’ It is noteworthy that stress shift does not apply to the comparative forms of adjectives, which are formed with *-um-* and retain the basic penultimate stress pattern of the language: *gogó* ‘strong’: *g<um>ógo* ‘stronger,’ *dengán* ‘good’: *d<um>éngan* ‘better,’ *roá* ‘bad’: *r<um>óa* ‘worse,’ etc.

In addition to the simple use of stress to mark morphological distinctions, some languages mark the vocative form of kinship terms with simultaneous stress shift and suffixation, as with Hanunóo (north-central Philippines) *ámaʔ* ‘father (ref.):’ *amá-ŋ* ‘father (voc.),’ *ínaʔ* ‘mother (ref.):’ *iná-ŋ* ‘mother (voc.).’

30.2.5 Compounding

Most of the languages from which data has been cited so far are found in Taiwan, the Philippines, and western Indonesia. As one moves into eastern Indonesia and the Pacific, however, the AN languages become morphologically much more impoverished (especially in parts of eastern Indonesia). With few affixes to use, word formation in these languages must necessarily draw more heavily on the combination of free morphemes, hence compounding.

Crowley (1982: 87) describes compounding as ‘a very productive noun-deriving process’ in the Paamese of central Vanuatu. He notes that such constructions are words rather than phrases since (i) their constituents cannot be separated by any modifier, and (ii) “it is not possible to left-dislocate any of the constituent stems of a compound as it is with any complex nominal phrase.” Among examples that he cites are *aute netano* ‘world’ (lit. ‘place down’), *avoŋi elua* ‘Tuesday’ (lit. ‘day two’), and *riou holauai* ‘coconut crab’ (lit. ‘hermit crab climb’).

In addition to such language-specific or area-specific patterns of compounding, some compounds with cognate elements are so widely distributed in AN languages that they must have been present in at least PMP. A number of these are words formed by joining a reflex of PMP **anak* ‘child, offspring; component part’ with another morpheme, as

in Malay *anak bulan* (a. + moon) ‘person of uncertain temper,’ *anak dagaŋ* (a. + trade, commerce) ‘foreigner,’ *anak dara* (a. + virgin) ‘maiden,’ *anak kapal* (a. + ship) ‘crew of a ship,’ *anak kunci* (a. + lock) ‘key of a lock,’ *anak sunai* (a. + river) ‘tributary stream,’ *anak tangga* (a. + ladder) ‘rung of a ladder,’ or the Manggarai (Flores, Lesser Sunda islands) terms *anak wina* (a. + female) ‘wife-takers,’ *anak rona* (a. + male) ‘wife-givers’ (the apparent reversal of gender here is due to the native conceptualization of the wife as coming from the symbolically “male” = superior group, which provides the source of life to the wife-takers, and hence is its material and spiritual benefactor).

Undoubtedly the most richly-developed source of compounds is the set of words that incorporate reflexes of PMP *mata ‘eye’ as part of a figurative expression. Malay, again, is a source of many examples, as with *mata aŋin* (m. + wind) ‘point of the compass,’ *mata ayər* (m. + fresh water) ‘spring of water,’ *mata hati* (m. + liver) ‘mental perception, instinct,’ *mata ikan* (m. + fish) ‘callus, corn, wart,’ *mata kaki* (m. + foot) ‘ankle,’ *mata pisau* (m. + knife) ‘blade of a knife,’ or the well known *matahari* (generally written as one word instead of two, but clearly m. + day) ‘sun.’ Although these appear to be best analyzed as nominal compounds in Malay, the cognate forms in other languages show that they probably originated as genitive phrases, as seen with Fijian *mata ni caŋi* (m. + GEN + wind) ‘point of the compass,’ *mata ni wai* (m. + GEN + fresh water) ‘spring of water,’ or *mata ni ika* (m. + GEN + fish) ‘sore on foot or hand.’

30.3 AN IN-DEPTH LOOK AT ONE WORD-FORMATION PROCESS: REDUPLICATION

In this final section I will look at one type of word-formation process that takes many different forms, namely reduplication.

Probably most of the phonological patterns and grammatical functions associated with reduplication in the world’s languages are found in AN languages, where reduplication tends to play a much larger role than in Indo-European, or many other languages. The following is a description, in as much depth as I can manage within the scope of this chapter, of the types of patterns and associated functions that have been reported.

30.3.1 Full Reduplication

Full reduplication in AN languages takes two forms: complete copying of the base, and complete copying of the base minus the coda. The latter pattern, which I will address first is found in many languages in Taiwan, the Philippines, and western Indonesia, of which Thao (central Taiwan) can serve as an example: *cpiq* ‘whip, beat’: *cpi-cpiq* ‘whip repeatedly,’ *kirac* ‘light, luminosity’: *pish-kira-kirac* ‘send out sparks,’ *lhun* ‘wet nasal

mucus': *mak-lhu-lhun* 'blow the nose repeatedly,' Itbayaten (northern Philippines) *hapin* 'mat': *h<in>api-hapin* 'what has been woven,' *koxat* 'heat': *ma-ŋoxat* 'to boil,' *ma-ŋoxa-ŋoxat* 'boil repeatedly,' Ngaju Dayak (southeast Borneo) *abas* 'strong': *aba-abas* 'rather strong,' *humon* 'stupid': *humo-humon* 'rather stupid,' or Bolaang Mongondow (northern Sulawesi) *bayag* 'bright': *ko-baya-bayag-an* 'become completely bright or clear,' *posad* 'communal work': *ko-posa-posad* 'do communal work.' Like complete copying of the base, full reduplication minus the coda generally signals repeated action with dynamic verbs, and often signals diminution in size for nouns, or decrease of intensity for adjectives/stative verbs (see Sections 30.3.2 and 30.3.6).³

Complete copying of the base is seen, among other languages, in Malay and Karo Batak of western Indonesia, where it generally marks plurality of nouns, and repeated action with verbs, as in Malay *anak* 'child': *anak-anak* 'children,' *rumah* 'house': *rumah-rumah* 'houses,' Karo Batak *tulan* 'bone': *tulan-tulan* 'bones,' or Malay *mə-lihat* 'to see': *mə-lihat-lihat* 'to look around (as at a museum or in shopping),' Karo Batak *bual* 'boast, lie': *ər-bual-bual* 'to tell tall stories,' *daram* 'seek': *tər-daram-daram* 'looking here and there.' With intransitive verbs full reduplication in Malay sometimes signals the leisurely execution of an action, as with *bər-jalan* 'to walk': *bər-jalan-jalan* 'take a leisurely stroll,' or *bər-pikir* 'to think': *bər-pikir-pikir* 'think over, consider, weigh the advantages of taking some action,' where it is ambiguous for repetition or leisurely execution.

In a number of languages in the Philippines and western Indonesia full reduplication with or without additional affixation marks simulative nouns, as in Kapampangan (central Luzon) *anák* 'child': *anak-anák-an* 'adopted child,' Botolan Sambal (west-central Luzon), *anak* 'child': *anak-anak* 'doll,' *bali* 'house': *bali-bali* 'playhouse,' tawo 'person': *tawo-tawo* 'scarecrow,' or Old Javanese *anak* 'child': *anak-anak-an* 'doll.' In Bikol (southeast Luzon) a similar pattern signals a related semantic property with verbs, namely feigning an action, as with *búta* 'blind': *búta-búta* 'feigning blindness,' or *buñóg* 'deaf': *buñóg-buñóg* 'feigning deafness' (Mintz 1971: 150).

Since a fully repeated disyllable is ambiguous for full or suffixal foot reduplication, it is important to ask whether polysyllabic bases also show complete copying, and this appears to be the case both in Malay (*perpustakaan* 'library': *perpustakaan-perpustakaan* 'libraries') and in Karo Batak (*sinuan* 'a plant': *sinuan-sinuan* 'plants').

³ 'Full reduplication minus the coda' may sound like a contradiction in terms, but it is difficult to see how else to characterize this pattern. If one were to argue that these are examples of CVCV reduplication we would expect the reduplicated form of trisyllables with a final consonant such as *quliush* 'long' to be ***quli-quliush*, but the actual form is seen in *mia-quliu-liush* 'to straighten, stretch out.' Likewise, bases of the shape CVCCVC cannot be characterized as exhibiting a pattern of CVCCV—reduplication, as shown by *buqnur* 'anger, hatred,' but *mia-buqnur-qnur* (not ***mia-buqnur-buqnur*) 'to be irritable.' The only clear generalization from such data is that patterns of both full and suffixal reduplication disallow the coda of the base to occur in surface forms.

Table 30.2 The reduplication–transitivity correlation in Tok Pisin

Transitive	Intransitive
wasim 'to wash s.t.'	waswas 'to bathe'
tingim 'to remember, think of'	tingting 'to think, ponder'
lukim 'to see'	lukluk 'to look'
tokim 'to say s.t., speak to'	toktok 'to talk'

In Thao, however, what appears to be full reduplication with disyllabic bases turns out on closer inspection to be suffixal foot reduplication, as in *kikalhi* 'to ask': *ma-kikalhi-kalhi* 'to ask around,' *shnara* 'ignite, catch fire': *pa-shnara-nara* 'burn something repeatedly,' or *qriu?* 'to steal': *q<un>riu-riu?* 'steal repeatedly.' Although the matter cannot be taken up in detail here, the recognition that Thao *cpi-cpiq* 'whip repeatedly,' and similar patterns, are instances of suffixal reduplication rather than full reduplication minus the coda carries the interesting and controversial theoretical implication that the reduplicant preserves more phonemic material from the underlying form than is found in the surface form of the base (for further discussion see fn. 3, and Blust 2009: 407–11).

Thieberger (2006: 194ff) reports a number of distinct grammatical functions for full reduplication of verb stems in the South Efate language of Vanuatu. These include: (i) diminution (*pes* 'speak': *fesfes* 'cheep (of a bird),' *lom* 'wet': *lomlom* 'moist'); (ii) iteration or intensification (*fek* 'to show': *fekfek* 'to show off,' *mat* 'to die': *matmat* 'to die (of many),' *pul* 'to twirl': *fulful* 'to spin'); (iii) nominalization (*si* 'to blow/shoot': *sisi* 'rifle,' *sif* 'to sling': *sifsif* 'slingshot'); and (iv) detransitivization (*fis* 'to whip' (trans.): *fis-fis* 'to masturbate' (intrans.), *wes* 'to work' (trans.): *weswes* 'to work' (intrans.), *si* 'to blow (trans.): *sisi* 'to blow' (intrans.)). The latter pattern is in fact widespread in Oceanic languages, and has even been incorporated in the English-derived lexicon of Tok Pisin, the national language of Papua New Guinea, as seen in Table 30.2.

30.3.2. CV- Reduplication

CV- reduplication is a common pattern in AN languages, tending to occur much more with verbs or adjectives than with nouns. In Bunun (central Taiwan) it marks durative aspect, collectivity, or intensity, in Tagalog it marks future tense, and in Manam of New Guinea it marks plural adjectives: Bunun *ma-asik* 'to sweep': *ma-a-asik* 'keep sweeping,' *ma-patas* 'to write': *ma-pa-patas* 'keep writing,' *ma-bulav* 'ripe': *ma-bu-bulav* 'all ripening at once,' *ma-kuis* 'slender': *ma-ku-kuis* 'very slender'; Tagalog *b<um>ilí* 'to buy': *bi-bilí* 'will buy,' *l<um>ákad* 'to walk': *la-lákad* 'will walk'; Manam (Lichtenberk 1983) *salaga* 'long': *sa-salaga* 'long (pl.),' *tumura* 'cold': *tu-tumura* 'cold (pl.),' *no?a* 'ripe': *no-no?a* 'ripe (pl.)'. A larger sample of languages undoubtedly would turn up additional functions associated with this pattern.

Although CV- reduplication is usually described as an unambiguous pattern, it typically targets an initial consonant slot (which may be vacuous) and the first vowel, as in Bunun *ma-a-asik* ‘keep sweeping,’ Pangasinan (north-central Luzon) instead copies CV, where C is the first consonant of the word, whether it is initial or not, as in *amígo* ‘friend’: *a<mi>mígo* (***a-amígo*) ‘friends,’ *kanáyon* ‘relative’: *ka-kanáyon* ‘relatives,’ *nióg* ‘coconut’: *ni-nióg* ‘coconuts,’ etc.

30.3.3 CVC- Reduplication

This pattern, sometimes called ‘heavy syllable reduplication’ (Hayes and Abad 1989), copies a syllable when the copied portion precedes a consonant in the base, but a syllable plus onset when the copied portion precedes a vowel, as in Agta (northern Luzon) *uf.fu* ‘thigh’: *u.f-uf.fu* ‘thighs,’ *tak.ki* ‘leg’: *tak-tak.ki* ‘legs,’ where it marks plural nouns, in *d<u.m>a.taŋ* ‘arrive’: *d<u.m>at.-da.taŋ* ‘keep arriving (one after the other),’ where it indicates repetitive action with at least intransitive verbs, in *a.pí.si* ‘small’: *a.p-a.pí.si* ‘very small,’ *ma.-ba.ŋi* ‘delicious-smelling’: *ma.-baŋ.ba.ŋi* ‘very delicious-smelling,’ where it marks intensity, and in *á.tu* ‘dog’: *a.t-á.tu* ‘puppy,’ or *ba.lá.taŋ* ‘girl’: *bal.-ba.lá.taŋ* ‘little girl,’ where it marks diminutive with some nouns.

30.3.4 Ca- Reduplication

This pattern, which copies the base-initial consonant followed by the fixed vowel /a/ (which is the full reduplicant in vowel-initial bases), is found in a number of Formosan and Malayo-Polynesian languages, where it is used to form instrumental nouns, and to mark [+human] numerals. Asilulu, of the central Moluccas, can be used to illustrate the first function, and Motu, of Southeast New Guinea, the second: Asilulu *heu* ‘use levers to raise a boat’: *ha-heu* ‘levers used to raise a boat,’ *lima* ‘to bail water’: *la-lima wael* ‘bailing well in boat,’ *sua* ‘to pry out’: *sa-sua niri* ‘toothpick,’ *tohi* ‘to strain (foods)’: *ta-tohi* ‘a strainer’; Motu *rua* ‘two’: *ra-rua* ‘two (people),’ *toi* ‘three’: *ta-toi* ‘three (people),’ *hitu* ‘seven’: *ha-hitu* ‘seven (people).’ Where the first base vowel is /a/ this pattern is formally identical to CV- reduplication, although the semantic patterning generally makes the distinction between these types clear, as in Asilulu *hala* ‘carry on the shoulder’: *ha-hala* ‘carrying pole’ (Ca-, not CV-), or Motu *hani* ‘four’: *ha-hani* ‘four (people).’

30.3.5 Infixal Reduplication

Reduplicative infixes have been reported in only a few AN languages. Among these is Thao, which uses Ca- reduplication both as a prefix and as an infix. Examples of the latter are seen in *pashʔuzu* ‘phlegm’: *mash<ʔa>ʔuzu-ʔuzu* ‘cough repeatedly, as in

trying to bring up phlegm' (Ca copying occurs within the base), and in *pakin-tutuz* 'to stack up': *p<in>a<ka>kin-tutuz-in* 'was stacked up by someone' (Ca copying occurs within the prefix *pakin-*).

30.3.6 Suffixal Reduplication

Suffixal foot reduplication has already been described in connection with patterns of full, or apparently full reduplication. However, suffixal syllable reduplication is also found in Chamorro, where it marks intensity in adjectives, and in several Oceanic languages, where its function is more difficult to pinpoint: Chamorro (Topping 1973) *ñalan* 'hungry': *ñala-lan* 'very hungry', *metgot* 'strong': *met-go-got* 'very strong' (with coda-dropping in the base, as with full reduplication in many languages), *bunita* 'pretty': *bunita-ta* 'very pretty', Sye (southern Vanuatu; Crowley 1998) *acumsu* 'black': *acumsu-su* 'pitch black', *ourup* 'to lob': *ouruv-rup* 'to lob', *owatop* 'walk to meet oncoming person': *owatov-top* 'walk stealthily in search of prey'.⁴

30.3.7 Odd Reduplicative Patterns

Space hardly permits an adequate survey of the full range of patterns and functions of reduplication in AN languages, but some patterns are typologically quite unusual, as with the atypical CVC- reduplicant of Palawano in the west-central Philippines (Revel-Macdonald 1979: 187), which copies the initial CV and *final* consonant of a nominal base to signal diminutive, and sometimes other shades of meaning: *kusiŋ* 'cat': *kun-kusiŋ* 'kitten', *baju?* 'clothing': *baʔ-baju?* 'child's clothing', *libun* 'woman': *lin-libun* 'girl', *kunit* 'yellow': *kut-kunit* 'yellow flycatcher (bird)'. Other unusual patterns of reduplication feature exceptionally long stretches of fixed segmentism, as in Agta of northern Luzon, where *Cala-* marks diminutive nouns (Healey 1960): *báhuy* 'pig': *bala-báhuy* 'a little pig', *pirák* 'money': *pala-pirák* 'a little money', *wer* 'creek': *wala-wer* 'a small creek'.

30.3.8 Triplication

The last mechanism of word formation in AN languages that I will consider is triplication—a device that has rarely been reported in the world's languages (Inkelas and Zoll 2005). Thao, of central Taiwan, can again be used to illustrate. In this language

⁴ An alternative interpretation would be to analyze these forms as containing reduplicative infixes. However, apart from some cases of *Ca* reduplication, all known infixes in AN languages (both independent and reduplicative) have the shape -VC-, and never immediately precede the final syllable of a polysyllabic word (as Chamorro *bunita*, or Sye *owatop*). General considerations of canonical shape in polymorphemic words thus disfavor this analysis.

triplication is formally expressed in either of two ways: (i) by a single pattern of reduplication used iteratively, or (ii) by two different patterns of reduplication used in combination. Examples of the first type are seen in *m-apa* 'to carry': *apa-apa-apa-n* [apapápan] 'be carried,' or *shkash* 'fear': *makit-shka-shka-shkash* 'slowly overcome by a sense of apprehension or foreboding' (with full reduplication applied twice) or *tapʔan* 'a patch on clothing': *t<in>apʔa-pʔa-pan* 'was patched repeatedly, of tattered clothing' (with suffixal foot reduplication applied twice, and automatic coda reduction). An example of the second type was seen in *pashʔuzu* 'phlegm': *mash<ʔa>ʔuzu-ʔuzu* 'cough repeatedly, as in trying to hawk up phlegm,' where both infixal Ca reduplication and suffixal foot reduplication combine (along with verbalizing pseudo nasal substitution triggered by the actor voice infix *-um-*) to signify repetition of the action.

CHAPTER 31

NIGER-CONGO

DENIS CREISSELS

31.1 INTRODUCTION

THE name Niger-Congo was introduced in 1955 by Joseph H. Greenberg, but is now used for the putative language phylum designated by Greenberg himself as Niger-Kordofan—see Greenberg (1955) and Greenberg (1963c). In most recent accounts, the Niger-Congo phylum is commonly presented as a language family with the following branches: Mande, Kordofanian, Atlantic, Ijoid, Kru, Kwa, Benue-Congo (including Bantu), Dogon, Gur, Adamawa, and Ubangian, with a total of approximately 1,500 languages. The validity of this grouping remains controversial however. As discussed in detail by Dimmendaal (2008, 2011), there is convincing evidence that a subset of the language groups enumerated above do constitute a genetically valid grouping for which the name Niger-Congo can be used, but serious doubts exist about the Niger-Congo affiliation of some others. The evidence supporting a genetic relationship between Benue-Congo, Kwa, Gur, and Adamawa is particularly solid, and these language groups can be viewed as constituting the core of the Niger-Congo phylum. By contrast, the inclusion of Mande, Ijoid, Dogon, and Ubangian into Niger-Congo is particularly problematic, and many specialists accept the view that it is safer to consider them as independent language families whose Niger-Congo affiliation cannot be considered as established on the basis of the available evidence.

This question has a direct impact on the organization of this chapter, since the main reason for setting apart Mande, Ijoid, Dogon, and Ubangian is that the language groups that constitute the core of the Niger-Congo phylum show particularly clear morphological evidence supporting the hypothesis of a genetic relationship in two subsystems that have to do with derivation: noun classification, and so-called verb extensions, whereas in Mande, Ijoid, Dogon, and Ubangian, reflexes of the Niger-Congo noun classes seem to be lacking, and the existence of verb suffixes cognate with the Niger-Congo verb extensions is problematic.

This means that the ancestors of the language groups that constitute the core of Niger-Congo shared a particular type of derivation system and that, in a historical

perspective, the departures from this prototype observed in present-day languages are innovations. Particularly conservative systems are common among Bantu and Atlantic languages.

For lack of space, this chapter is limited to a presentation of derivational processes involving the major lexical categories (nouns and verbs) in Niger-Congo languages whose derivation system is close to the Niger-Congo prototype. There is no real unity in the derivation systems of the languages in which the Niger-Congo system of noun classes and verb extensions has been more or less eroded. The only generalization that can be proposed is that the loss of the Niger-Congo derivational devices that will be presented in the following sections tends to be compensated by a development of compounding (which is rather marginal in typical Niger-Congo systems), and also, in the case of $V \rightarrow V$ derivation devices, by the development of serialization. This however does not imply any uniformity in the word formation systems of the languages in question, since subsequent grammaticalization phenomena may have resulted in the creation of various functional types of derivational processes whose detailed presentation would require much more space than is available here. As regards Mande, Ijoid, Dogon, and Ubangian, these language families show derivation systems that are not particularly close to the Niger-Congo prototype, and would therefore require separate presentations.

A very general characteristic of Niger-Congo languages with derivational systems close to the Niger-Congo prototype is the marginality (or even inexistence) of noun-to-verb derivation. This explains why the rest of this chapter is organized into three sections only, devoted to verb-to-verb derivation (Section 31.2), noun-to-noun derivation (Section 31.3), and verb-to-noun derivation (Section 31.4).

31.2 VERB-TO-VERB DERIVATION

Niger-Congo languages typically have suffixes known as “verb extensions” used to derive verbs from verbs with a variety of functions: verb extensions may increase the valency, decrease the valency, (re-)orient the action, or introduce aspectual specifications. In addition to the use of verb extensions, reduplication is widely attested, most often with an iterative meaning.

In some cases, the stems to which extensions attach are not attested, and the justification for segmenting the extension is the possibility of substituting other extensions, or the fact that the same ending is found in a series of verb stems sharing a common element of meaning. For example, Tswana has no verb **apa*, but comparing *apara*¹ ‘dress

¹ In most Bantu languages, verb forms end with a vowel that must be analyzed as an inflectional ending, since its variations contribute to the identification of the tenses that constitute the verbal paradigm. In particular, Tswana verbs are quoted in a form (the infinitive) that must be segmented as “verb stem + final vowel *a*.” For example, the stem of the verb quoted as *apara* is *apar-*, and the stem of *apola* is *apol-*.

(oneself), put on top garments' with *apola* 'undress, take off garments' makes it possible to identify in these two verbs an abstract stem *ap-* combined with the contactive extension *-ar-* and the separative extension *-ol-*.

As discussed in more detail in Hyman (2007), the following verb extensions have been reconstructed for Proto-Bantu² and have also been proposed for a reconstruction at Proto-Niger-Congo level by Voeltz (1977): applicative, causative, contactive/tentive, passive, reciprocal, reversive/separative (transitive), reversive/separative (intransitive), stative/neuter, stative/positional.

The applicative extension increases the valency by licensing an additional term in the syntactic role of object referring to a participant that otherwise could only be encoded as an adjunct, or could not be mentioned at all. The objects licensed by applicative derivation (commonly termed *applied objects*) may have a variety of semantic roles, but the use of applicative derivation to encode beneficiaries, illustrated in (1) by Tswana *apaya* 'cook' / *apeela* 'cook for someone', is particularly common. The same applicative suffix *-el-*, more or less modified by morphophonological processes, is further illustrated in examples (2)–(12) (in which the applied object is boxed) in a variety of other uses.³

- (1) a. *Lorato* *o-apay-a* *bo-gobe*.
 CL1.Lorato CL1-cook-FV CL14-porridge
 'Lorato is cooking porridge.'
- b. *Lorato* *o-ape-el-a* *b-ana* *bo-gobe*.
 CL1.Lorato CL1-cook-APPL-FV CL2-child CL14- porridge
 'Lorato is cooking porridge for the children.'
- (2) a. *Kgosi* *e-atlhol-ets-e* *mo-nna* *bo-godu*.
 CL9.king CL9-condemn-APPL.PRF-FV CL1-man CL14-theft
 'The king condemned the man for theft.'
- b. *Kgosi* *e-atlhol-ets-e* *mo-nna* *lo-so*.
 CL9.king CL9-condemn-APPL.PRF-FV CL1-man CL11-death
 'The king condemned the man to death.'
- (3) *Kitso* *o-berek-el-a* *tiego*.
 CL1.Kitso CL1-work-APPL-FV CL9.delay
 'Kitso is working in order to make up lost time.'
- (4) *Ma-godu* *a-bola-ets-e* *mo-nna* *ma-di*.
 CL6-thief CL6-kill-APPL.PRF-FV CL1-man CL6-money
 'The thieves killed the man for money.'

² For a detailed presentation of the derivational processes reconstructed for Proto-Bantu (in particular, verb extensions), see Schadeberg (2003).

³ The following abbreviations are used in the glosses of the Tswana examples: APPL = applicative, CLX (X a number between 1 and 17) = noun class marker of class X, DEM = demonstrative, FV = final vowel, GEN = genitive, PASS = passive, PRF = perfect, SG = singular.

- (5) *Mpho o-j-ets-e* Kitso *di-nawa.*
 CL1.Mpho CL1-eat-APPL.PRF-FV CL1.Kitso CL10-bean
 ‘Mpho ate the beans to Kitso’s detriment.’
 (to be understood as ‘Mpho ate the beans that were intended for Kitso.’)
- (6) *Ke-lebog-el-a* *Kitso* ma-di.
 1SG-thank-APPL-FV CL1.Kitso CL6-money
 ‘I am thanking Kitso for the money.’
- (7) *Mo-sadi yo o-ak-el-a* ralebentlele.
 CL1-woman CL1.DEM CL1-lie-APPL-FV CL1.shopkeeper
 ‘This woman is telling lies about the shopkeeper.’
- (8) *Ma-bele a-alaf-el-w-a* tshupa.
 CL6-sorghum CL6-treat-APPL-PASS-FV CL9.tshupa
 ‘The sorghum is treated against *tshupa* (a kind of worm).’
- (9) a. *Mo-sadi o-bil-ets-a* *b-ana* di-jo.
 CL1-woman CL1-call-APPL-FV CL2-child CL10-food
 ‘The woman is calling the children to eat (lit. ‘for food’).’
 b. *Mo-sadi o-bil-ets-a* b-ana *ngaka.*
 CL1-woman CL1-call-APPL-FV CL2-child CL9.doctor
 ‘The woman is calling the doctor for the children.’
- (10) a. *Ke-f-ets-e* *ngwana-ke* baesekele *ma-di.*
 1SG-give-APPL.PRF-FV CL1.child-1SG CL9.bicycle CL6-money
 ‘I gave money to my son for a bicycle.’
 b. *Ke-f-ets-e* malom-e *di-kgomo le-tswai.*
 1SG-give-APPL.PRF-FV CL1.uncle-1SG CL10.cow CL5-salt
 ‘I gave salt to the cows for my uncle.’
- (11) *Mo-setsana yo o-fos-ets-a* ba-tsadi.
 CL1-girl CL1.DEM CL1-be_wrong-APPL-FV CL2-parent
 ‘This girl behaves badly towards her parents.’
- (12) *Ke-gakgamal-el-a* bo-pelokgale jw-a-ngw-ana yo.
 1SG-be_impressed-APPL-FV CL14-courage 14-GEN-1-child CL1.DEM
 ‘I am impressed by the courage of this child.’

The causative extension increases the valency by introducing a causer in subject role, whereas the subject of the non-derived verb is demoted to object, as in Tswana *opela* ‘sing’ / *opedisa* ‘cause to sing, conduct (a choir).’

The contactive (or tentative) extension is a non-productive extension found in verbs that have in common the meaning of actively marking firm contact, as in Tswana *apara* ‘put on top garments’ (previously mentioned in this section).

The passive extension decreases the valency by demoting the subject and promoting the object in subject role, as in Tswana *roma* ‘send’ / *rongwa* ‘be sent.’

The reciprocal extension, in addition to the reciprocal meaning it expresses with semantically bivalent verbs (as in Tswana *rata* ‘love’ / *ratana* ‘love one another’), is often

found with an associative meaning ('do s.t. together,' as in Tswana *bopega* 'take shape' / *bopagana* 'fuse') in combination with monovalent verbs. It may also express repetitive actions, which is reminiscent of the range of meanings characteristic of so-called "pluractional" markers. Among Bantu languages, antipassive uses of this extension are common, as in Rundi *tuka* (transitive) 'insult' / *tukana* (intransitive) 'insult one another' (reciprocal) or 'insult people' (antipassive). In some Bantu languages, the reciprocal use of this extension tends to become obsolete, and the reciprocal function tends to be taken over by the reflexive object index or by the so-called "neuter" extension (anticausative).

The reversive (or separative) extensions imply movement out of some original position, as in Swahili *ziba* 'block' / *zibua* 'unblock.' Bantu languages typically have two distinct separative extensions for spontaneous movement (intransitive) and caused movement (transitive), as in Tswana *ama* 'touch' / *amologa* 'become separated' / *amolola* 'separate.'

The so-called "stative" or "neuter" extension converts transitive verbs into intransitive verbs expressing an anti-causative meaning, as in Tswana *senya* 'spoil' / *senyega* 'become spoilt' or *bona* 'see' / *bonala* 'be visible, appear.'

The "stative/positional" extension is found in verbs expressing "assume a position, be in a position," as in Tswana *botha* 'repose' / *bothama* 'repose comfortably, lie at ease.'

The "impositive" extension, which is not mentioned in the list above but is reconstructed for Proto-Bantu, is functionally a variety of causative involving direct causation as well as a locative element of meaning ("put something into some position"); it constitutes the transitive counterpart of the "stative/positional" extension, as in Tswana *sekama* 'lean against (intr.)' / *sekega* 'lean against (tr.).'

Among the verb extensions that do not modify valency, reversive (or separative) is particularly common. For example, in Jóola Banjal (Atlantic), *eppég-ul* 'open' derives from *eppek* 'shut' in the same way as *efóg-ul* 'dig up' from *efok* 'bury' Bassène (2007: 69). Itive / ventive extensions expressing centripetal / centrifugal movement (as in Wolof *sàcc* 'steal' / *sàcc-i* 'go and steal' / *sàcc-si* 'come and steal'), although not mentioned in the list above, are also relatively common.

To give an idea of the variety observed in the inventories of verb extensions, in the rest of this section, I briefly present the verb suffixes used with a valency-changing function in Wolof (Atlantic). The data are taken from Nouguié Voisin (2002).

In Wolof, the valency changes systematically coded by means of verb suffixes can be classified into six types: middle, causative, applicative, co-participation (including reciprocal), antipassive, and possessive. A puzzling feature of the coding of valency changes in Wolof is that, as shown by Table 31.1, similar valency changes may be coded by different suffixes, and the same suffix may code different valency changes, also summarized in Table 31.1.

Note that Wolof has no passive extension. Some uses of the middle marker *-u* can be considered as quasi-passive, but constructions combining object topicalization and subject focalization are more commonly used in Wolof with a function similar to that fulfilled by passive constructions in other languages.

Table 31.1 Valency change types and valency change markers in Wolof

type of valency change	possible markers
middle	-u
causative	-e, -a/, -le, -lu, -loo
applicative	-e, -a/
co-participation	-e, -oo, -ante, -andoo
antipassive	-e
possessive	-le

Middle *-u*, illustrated by (13)–(15), is typically used to express a reflexive or anticausative meaning.⁴

- (13) a. *Jigéen j-i sang na xale b-i.*
 woman CL-DEF wash PRF.3SG child CL-DEF
 ‘The woman washed the child.’
 b. *Xale b-i sang-u na.*
 child CL-DEF wash-MID PRF.3SG
 ‘The child washed.’
- (14) a. *Těj naa bunt b-i.*
 close PRF.1SG door CL-DEF
 ‘I closed the door.’
 b. *Bunt-u néeg b-i du téej-u b-u baax.*
 door-H room CL-DEF NEG.EMPH.3SG close-MID CL-LNK be_good
 ‘The door of the room does not close well.’
- (15) a. *Omar daan na Usmaan.*
 Omar bring-down PRF-3SG Ousmane
 ‘Omar brought Ousmane down.’ ‘Omar defeated Ousmane.’
 b. *Garab g-i daan-u na.*
 tree CL-DEF bring-down-MID PRF.3SG
 ‘The tree fell down.’

⁴ The following abbreviations are used in the glosses of Wolof examples: ANTIP = antipassive, APPL = applicative, CAUS = causative, CL = noun class marker, DEF = definite, DEM = demonstrative, DIST = distal, EMPH = emphatic, FOC = focalization, H = head-marking, INDEF = indefinite, LNK = attributive linker, MID = middle, NEG = negation, PL = plural, PRF = perfect, PROX = proximal, RECP = reciprocal, REL = relativizer, SG = singular, SOC = sociative.

Causativization can be encoded by five different suffixes as illustrated in (16)–(20). They differ in semantic nuances. The clearest case is that of *-le*, specialized in the expression of sociative causation.

- (16) a. *Usmaan genn na.*
 Ousmane go-out PRF₃SG
 ‘Ousmane went out.’
 b. *Usmaan genn-e na woto b-i.*
 Ousmane go-out-CAUS PRF₃SG car CL-DEF
 ‘Ousmane got out the car.’
- (17) a. *Xale b-i toog na.*
 child CL-DEF sit-down PRF₃SG
 ‘The child sat down.’
 b. *Jigéen j-i toog-al na xale b-i.*
 woman CL-DEF sit-down-CAUS PRF₃SG child CL-DEF
 ‘The woman sat the child down.’
- (18) a. *Těj naa bunt b-i.*
 close PRF₁SG door CL-DEF
 ‘I closed the door.’
 b. *Těj-lu naa bunt b-i.*
 close-CAUS PRF₁SG door CL-DEF
 ‘I made (someone) close the door.’
- (19) a. *Góor g-i toog na.*
 man CL-DEF sit-down PRF₃SG
 ‘The man sat down.’
 b. *Toog-loo nañu góor g-i.*
 sit-down-CAUS PRF₃PL man CL-DEF
 ‘They made the man sit down.’
- (20) a. *Xale b-i lekk na.*
 child CL-DEF eat PRF₃SG
 ‘The child ate.’
 b. *Jigéen j-i lekk-le na xale b-i.*
 woman CL-DEF eat-SOC.CAUS PRF₃SG child CL-DEF
 ‘The woman helped the child eat.’

As illustrated above by Tswana, languages with a single multifunctional applicative marker are very common in Niger-Congo, in particular among Bantu languages. By contrast, Wolof has two relatively specialized applicative extensions: *-al* encodes the addition of an object representing a beneficiary, as in (21), or the conversion of a comitative adjunct into an object, as in (22), whereas the objects licensed by the applicative suffix *-e* refer to instruments, means, or places, as illustrated in (23)–(24).

- (21) a. *Rey nañu a-m xar.*
kill PRF₃PL INDEF-CL sheep
'They killed a sheep.'
- b. *Rey-al nañu gan g-i a-m xar.*
kill-APPL PRF₃PL guest Cl-Def Indef-Cl sheep
'They killed a sheep for the guest.'
- (22) a. *Usmaan séy na ak Faatu.*
Ousmane marry Prf₃sg with Fatou
'Ousmane married Fatou.'
- b. *Faatu la Usmaan séy-al.*
Fatou FOC₃SG Ousmane marry-APPL
'It is Fatou that Ousmane married.'
- (23) a. *Añ nañu ak ceeb-u jën.*
lunch PRF₃PL with rice-H fish
'They had rice and fish for lunch.'
- b. *Añ-e nañu ceeb-u jën.*
lunch-APPL PRF₃PL rice-H fish
same meaning as (a)
- (24) a. *Añ nañu c-i teraas b-i.*
lunch PRF₃PL LOC-PROX terrace CL-DEF
'They had lunch on the terrace.'
- b. *C-i teraas b-i lañu añ-e.*
LOC-PROX terrace CL-DEF FOC₃PL lunch-APPL
'It is on the terrace that hey had lunch.'

In contrast to languages that have a single reciprocal or reciprocal/associative (here again, a situation typically found among Bantu languages), Wolof has four verbal extensions encoding various types of co-participation: *-andoo* (sociative) (25), *-oo* (sociative/reciprocal) (26), *-e* (reciprocal) (27), and *-ante* (reciprocal) (28).

- (25) a. *Xale b-i toog na c-i lal b-i.*
child CL-DEF sit-down PRF₃SG LOC-PROX bed CL-DEF
'The child sat on the bed.'
- b. *Xale y-i toog-andoo nañu c-i lal b-i.*
child CL-DEF sit-down-SOC PRF₃PL LOC-PROX bed CL-DEF
'The children sat on the bed together.'
- (26) a. *Sédd naa ko c-i xaalis b-i.*
share PRF₁SG 3SG LOC-PROX money CL-DEF
'I gave him part of the money.'
- b. *Sédd-oo nanu xaalis b-i.*
share-SOC PRF₁PL money CL-DEF
'We shared the money.'

- (27) a. *Omar gis na Usmaan c-a marse b-a.*
 Omar see PRF₃SG Ousmane LOC-DIST market CL-DEF
 ‘Omar saw Ousmane at the market.’
 b. *Omar ak Usmaan gis-e nañu c-a marse b-a.*
 Omar with Ousmane see-RECP PRF₃SG LOC-DIST market CL-DEF
 ‘Omar and Ousmane met (saw each other) at the market.’
- (28) a. *Rey nañu a-m xar.*
 kill PRF₃PL INDEF-CL sheep
 ‘They killed a sheep.’
 b. *Rey-ante nañu.*
 kill-RECP PRF₃PL
 ‘They killed one another.’

Antipassive *-e* encodes the deletion of patients of some transitive verbs, as in (29), and recipients of ditransitive verbs (in particular, of ditransitive verbs created by applicative derivation, as in (30)).

- (29) a. *Xaj-a ko màtt.*
 dog-FOC₃SG ₃SG bite
 ‘A dog bit him.’
 b. *Xaj b-ii du màtt-e.*
 dog CL-DEM NEG.EMPH₃SG bite-ANTIP
 ‘This dog does not bite.’
- (30) a. *Togg naa yàpp w-i.*
 cook PRF₁SG meat CL-DEF
 ‘I cooked the meat.’
 b. *Togg-al naa la yàpp w-i.*
 cook-APPL PRF₁SG ₂SG meat CL-DEF
 ‘I cooked the meat for you.’
 c. *Togg-al-e naa yàpp w-i.*
 cook-APPL-ANTIP PRF₁SG meat CL-DEF
 ‘I cooked the meat for some people.’

“Possessive” *-le* is a cross-linguistically rare type of valency operator (it has so far only been found in a few Senegalese languages). It combines with verbs that express a quality attributed to their subject; the same quality is attributed to the object of the derived verb, whereas the subject of the derived verb takes the role of possessor, as in (31).

- (31) a. *Woto b-i gaaw na.*
 car CL-DEF be-fast PRF₃SG
 ‘The car is fast.’

b. <i>Gaaw-le</i>	<i>naa</i>	<i>woto.</i>
be-fast-POSS	PRE.1SG	car
'I have a fast car.'		

31.3 NOUN CLASS SYSTEMS AND NOUN-TO-NOUN DERIVATION

The so-called noun class systems found in Niger-Congo languages can be viewed as a particular type of gender systems. A brief characterization of noun class systems is in order here, since in the Niger-Congo languages that have a noun class system close to the Niger-Congo prototype, gender change is a very productive way to derive nouns from nouns, whereas noun compounding as well as the use of affixes to derive nouns from nouns is marginal.

The notion of noun class in the description of Niger-Congo languages refers to a division of nouns into subsets manifested in their behavior in agreement mechanisms. The illustrations in this section are taken from Tswana, a southern Bantu language whose noun classification system stands very close to the prototype.

The essential feature of Niger-Congo noun class systems is that

- (a) noun forms divide into subsets (noun classes) according to their behavior in agreement mechanisms observed in the formation of noun phrases by combining a head noun with various types of modifiers, in the use of pronouns, and in the indexation of arguments on the verb;
- (b) the forms involved in these agreement mechanisms (nouns, noun modifiers, pronouns, and verbs) include affixes (class markers) that determine their behavior as controllers or targets of agreement; and
- (c) the classification is a lexicalized classification of nouns, and not a classification of referents directly, in the sense that it is not possible to change the class of a noun in order to emphasize particular semantic features of the same referent, as typically observed in so-called classifier systems.⁵

Semantically, the most obvious function of noun class markers (but not the only one—see examples (34) to (37)) is the expression of number (singular vs. plural). An important feature of Niger-Congo classification systems in their most typical form is the absolute impossibility to isolate morphemes expressing number independently from gender.

⁵ In Niger-Congo noun class systems, a given nominal lexeme may be compatible with several class markers, but the variations express changes in the referential meaning (as in example (34)), which is an entirely different semantic operation.

Moreover, there is generally no one-to-one correspondence between agreement classes and class markers in the singular and in the plural. This is the reason why many descriptions of Niger-Congo class systems do not emphasize the possibility of dividing noun *lexemes* into *genders*, but rather start from a division of noun *forms* into *classes* in which the singular form and the plural form of a given noun are treated as two distinct units; in this approach, a gender may be subsequently defined as a couple of classes that include the singular and plural forms of the same lexemes. For example, in Tswana, *mosadi* pl. *basadi* ‘woman’ belongs to gender 1/2, which means that its singular form *mosadi* belongs to the agreement class labeled “1,” and its plural form *basadi*, to the agreement class labeled “2.”⁶

In the description of Niger-Congo noun class systems, the regularity of agreement generally makes it easy to establish the number of classes into which noun forms divide (i.e. the number of possible agreement patterns); by contrast, the idiosyncrasies shown by many nouns in the singular–plural correspondence and the variations observed in plural formation often make it very difficult to decide how many genders must be recognized, if genders are defined as sets of nominal lexemes with the same agreement properties in both the singular and the plural.

The examples in (32) illustrate the division of Tswana noun forms into twelve classes on the basis of the agreement between nouns and adjectives in the attributive construction,⁷ and (33) illustrates Tswana noun forms of classes 1, 5, 7, and 9 (or the corresponding plural, if meaning requires it) in combination with various types of modifiers (genitives, demonstratives, etc.).

(32)	a. cl. 1	<i>mosadi yo moša</i>	‘new woman’
	b. cl. 2	<i>basadi ba baša</i>	‘new women’
	c. cl. 3	<i>molemo o moša</i>	‘new medicine’
	d. cl. 4	<i>melemo e meša</i>	‘new medicines’
	e. cl. 5	<i>lesaka le leša</i>	‘new cattle kraal’
	f. cl. 6	<i>maraka a maša</i>	‘new cattle kraals’
	g. cl. 7	<i>sekolo se seša</i>	‘new school’
	h. cl. 8–10	<i>dikolo tse dintšha</i>	‘new schools’
		<i>dikgosi tse dintšha</i>	‘new chiefs’
	i. cl. 9	<i>kgosi e ntšha</i>	‘new chief’
	j. cl. 11	<i>lokwalo lo loša</i>	‘new book’

⁶ The numbering of classes used in Bantu studies is based on the correspondence with the reconstructed classes of Proto-Bantu; for example, in the description of Tswana, “cl. 8–10” refers to a class that historically results from the merger of Proto-Bantu classes 8 and 10, and the absence of classes 12 and 13 means that no Tswana class is the reflex of Proto-Bantu classes 12 and 13.

⁷ This construction involves an obligatory linker expressing class agreement too. Historically, this linker is the reflex of a former demonstrative that has lost its semantic content and has become a purely formal element of the construction.

k.	cl. 14	<i>bojang jo boša</i>		'new grass'
l.	cl. 15–17	<i>go lema mo goša</i>		'new way of cultivating'
- (33)
- | | | | | |
|----|--------------------------------|----------------------|--|-------------------------------|
| a. | <i>mosadi wa motse</i> | | | 'woman of the village' |
| | <i>lekau la motse</i> | | | 'boy of the village' |
| | <i>sefofu sa motse</i> | | | 'blind person of the village' |
| | <i>ngaka ya motse</i> | | | 'doctor of the village' |
| b. | <i>mosadi yole</i> | | | 'that woman' |
| | <i>lekau lele</i> | | | 'that boy' |
| | <i>sefofu sele</i> | | | 'that blind person' |
| | <i>ngaka ele</i> | | | 'that doctor' |
| c. | <i>mosadi ofe?</i> | | | 'which woman?' |
| | <i>lekau lefe?</i> | | | 'which boy?' |
| | <i>sefofu sefe?</i> | | | 'which blind person?' |
| | <i>ngaka efe?</i> | | | 'which doctor?' |
| d. | <i>mosadi osele</i> | | | 'another woman' |
| | <i>lekau lesele</i> | | | 'another boy' |
| | <i>sefofu sesele</i> | | | 'another blind person' |
| | <i>ngaka esele</i> | | | 'another doctor' |
| e. | <i>mosadi ope</i> | | | 'no woman' |
| | <i>lekau lepe</i> | | | 'no boy' |
| | <i>sefofu sepe</i> | | | 'no blind person' |
| | <i>ngaka epe</i> | | | 'no doctor' |
| f. | <i>basadi botlhe</i> | | | 'every woman' |
| | <i>makau otlhe</i> | | | 'every boy' |
| | <i>difofu tsotlhe</i> | | | 'every blind person' |
| | <i>dingaka tsotlhe</i> | | | 'every doctor' |
| g. | <i>mosadi mongwe</i> | | | 'one woman' |
| | <i>lekau lengwe</i> | | | 'one boy' |
| | <i>sefofu sengwe</i> | | | 'one blind person' |
| | <i>ngaka nngwe</i> | | | 'one doctor' |
| h. | <i>basadi ba le babedi</i> | | | 'two women' |
| | <i>makau a le mabedi</i> | | | 'two boys' |
| | <i>difofu di le pedi</i> | | | 'two blind persons' |
| | <i>dingaka di le pedi</i> | | | 'two doctors' |
| i. | <i>mo-sadi</i> | <i>yo</i> | <i>ke-mo-thus-its-e-ng</i> | <i>maabane</i> |
| | CL ₁ -woman | CL ₁ .LNK | ₁ SG-CL ₁ -help-PRF-FV-REL | yesterday |
| | 'the woman I helped yesterday' | | | |

<i>le-kau</i>	<i>le</i>	<i>ke-le-thus-its-e-ng</i>	<i>maabane</i>
CL ₅ -boy	CL ₅ .LNK	1SG-CL ₅ -help-PRF-FV-REL	yesterday
'the boy I helped yesterday'			
<i>se-fofu</i>	<i>se</i>	<i>ke-se-thus-its-e-ng</i>	<i>maabane</i>
CL ₇ -blind	CL ₇ .LNK	1SG-CL ₇ -help-PRF-FV-REL	yesterday
'the blind person I helped yesterday'			
<i>ngaka</i>	<i>e</i>	<i>ke-e-thus-its-e-ng</i>	<i>maabane</i>
CL ₉ .doctor	CL ₉ .LNK	1SG-CL ₉ -help-PRF-FV-REL	yesterday
'the doctor I helped yesterday'			

If inflection is defined as the part of morphology directly involved in syntactic rules, the class markers of nouns in Niger-Congo languages are unquestionably inflectional. However, noun class systems tend to blur the traditional inflection vs. derivation distinction in the sense that, in addition to number, commonly integrated into noun inflection, class alternations express meanings more commonly expressed by derivational morphology.

The observation of class alternations expressing "derivational" meanings reveals semantic motivations that are not immediately apparent in the assignment of individual non-derived noun stems to genders defined on the basis of the expression of number. Noun stems may be involved in multiple class alternations, with, in particular, the following meanings: tree vs. fruit, individual vs. collective, concrete vs. abstract, animate entity vs. behavior (people vs. language, etc.), diminutive, augmentative.

The examples in (34) show two Tswana noun stems whose combination with class prefixes is not limited to the expression of number. Such variations in class prefixes express semantic distinctions that, cross-linguistically, tend to be expressed by means of derivational morphology.

- (34) a. cl. 1 *mosadi* 'woman'
 cl. 2 *basadi* 'women'
 cl. 7 *sesadi* 'feminine behavior'
 cl. 9 *tshadi* 'group of women'
 cl. 11 *losadi* 'group of women'
 cl. 14 *bosadi* 'womanhood'
- b. cl. 3 *moretlwa* 'tree of the sp. moretlwa'
 cl. 2 *meretlwa* 'trees of the sp. moretlwa'
 cl. 9 *thetlwa* 'fruit of the moretlwa tree'
 cl. 10 *dithetlwa* 'fruits of the moretlwa tree'
 cl. 11 *loretlwa* 'thicket of moretlwa trees'

Examples (35) and (36) illustrate the regular use of class alternations to express typically derivational meanings in Tswana: *concrete human* (cl. 1) ~ *abstract* (cl. 14) and *tree* (cl. 3) ~ *fruit* (cl. 9).

(35)	<i>mosadi</i>	‘woman’	/	<i>bosadi</i>	‘womanhood’
	<i>monna</i>	‘man’	/	<i>bonna</i>	‘manhood’
	<i>moloi</i>	‘witch’	/	<i>boloi</i>	‘witchcraft’
(36)	<i>moretlwa</i>	‘moretlwa tree’	/	<i>thetlwa</i>	‘moretlwa fruit’
	<i>morula</i>	‘morula tree’	/	<i>thula</i>	‘morula fruit’
	<i>mmilo</i>	‘mmilo tree’	/	<i>pilo</i>	‘mmilo fruit’

Tswana has lost the ability to express diminutive and augmentative by means of class alternations. This very common use of class alternations is illustrated in (37) by Jóola Banjal (Atlantic) from Bassène (2007).

(37)		<i>e-joba</i> ‘dog’	pl. <i>su-joba</i>
	(dim.)	<i>ju-joba</i> ‘small dog’	pl. <i>mu-joba</i>
	(augm.)	<i>ga-joba</i> ‘big dog’	pl. <i>ba-joba</i>

31.4 VERB-TO-NOUN DERIVATION

In typical Niger-Congo languages, verb-to-noun derivation typically involves two elements: the addition of a derivational suffix, and the addition of class morphology manifesting the assignment of the derived noun to a particular class, as in Jóola Banjal *-ffañ* ‘close’ → *e-ffañ-um* (pl. *si-ffañ-um*) ‘key,’ where *-úm* is a derivational suffix used to derive nouns of instruments from verbs, and *e-* / *si-* are class markers.

It may also happen that no overt derivational element is present, and the deverbal noun is formed by the mere addition of class morphology to a verb stem, as in Jóola Banjal *-mbal* ‘fish (V)’ → *e-mbal* (pl. *si-mbal*) ‘fish-trap.’ In such cases, the class to which such nouns are assigned may be crucial for the identification of their meaning.

Note that, as illustrated by the Tswana examples in the next paragraph, the addition of class morphology to verb stems may be responsible for consonant alternations in verb-to-noun derivation, since in many Niger-Congo languages, the initial consonants of noun stems alternate depending on the class to which the noun forms belong.

Among the nouns resulting from verb-to-noun derivation processes, the following functional types are particularly widespread: action nouns (as in Tswana *lema* ‘plough’ / *tem-o* (cl. 9) ‘ploughing’), result nouns (as in Tswana *bitsa* ‘call’ / *pits-o* (cl. 9) ‘meeting’), agent nouns (as in Tswana *lema* ‘cultivate’ / *mo-lem-i* (cl. 1) ‘farmer’), instrument nouns (as in Tswana *apara* ‘dress (oneself)’ / *se-apar-o* (cl. 7) ‘garment’), place-of-action nouns (as in Tswana *tlhaba* ‘slaughter’ / *ma-tlhab-elo* (cl. 6) ‘abattoir’), and manner nouns (as in

Tswana *apara* ‘dress (oneself)’ / *mo-apar-o* (cl. 3) ‘way of dressing’). Note that, with the exception of agent nouns (which are generally formed in a way that cannot be used for other semantic types of deverbal nouns), the general tendency is that a given morphological formation may be used for more than one semantic type of deverbal noun, and conversely, variations may occur in the expression of a given semantic type of deverbal noun.

Many languages have infinitives similar to action nouns in some respects but different from them in some others. In typical Niger-Congo languages, infinitives are morphologically similar to action nouns in the sense that their formation involves the addition of class morphology to a verb stem, and the phrases they head control class agreement in the same way as canonical noun phrases occupying the same positions in the clause. They differ from action nouns in the sense that the phrases headed by infinitives may include typical verb dependents that cannot be included in canonical noun phrases, and may be found in syntactic positions in which canonical noun phrases cannot be found. There are, however, important cross-linguistic and even intra-linguistic variations in the nouniness vs. verbiness of infinitives.

CHAPTER 32

AFROASIATIC

ERIN SHAY

32.1 THE AFROASIATIC PHYLUM

THE Afroasiatic language phylum, the fourth largest language phylum in the world, is comprised of about 375 living languages spoken by about 300 million speakers (<www.ethnologue.com>). Afroasiatic languages are spoken throughout North and Central Africa, the Horn of Africa, the Arabian Peninsula, and Central Asia (Arabic). While some Afroasiatic languages are widely spoken and have been thoroughly described, a great many languages in the phylum are undescribed and are spoken by a shrinking population. Because there is much work to be done on describing and analyzing such languages, any claims about typological or other universals among Afroasiatic languages must remain tentative.

The Afroasiatic phylum is usually divided into six language families (in alphabetical order): Berber, Chadic, Cushitic, Egyptian, Omotic, and Semitic. This division is taken for granted in the present work. The languages of the Berber family, a relatively homogeneous group, are spoken in Morocco, Algeria, Mali, Niger, Libya, Tunisia, and Burkina Faso, in isolated areas in Mauritania, and in the Siwa oasis in Egypt. The Chadic family is the largest in the Afroasiatic phylum, with 140–60 languages spoken in northern Nigeria, southern Niger, northern Cameroon, and southern Chad. Hausa, the Chadic language with the greatest number of speakers, serves as a vehicular language throughout much of West and Central Africa. Languages of the Cushitic family, which include Oromo, Somali, Sidamic, Dullay, Afar, and others, are spoken in the Horn of Africa, Sudan, Tanzania, Egypt, and Kenya. The Egyptian family consists of Ancient Egyptian and its sole descendant, Coptic, once the liturgical language of the Christian church of Egypt. The Omotic languages, the least-studied branch of Afroasiatic (Hayward 2003), include Ari, Bench, Dime, Dizi, and Kafa and other languages spoken in areas of south-central and western Ethiopia. The most widely studied languages of the Afroasiatic phylum belong to the Semitic family, which includes Arabic, Hebrew, Amharic, Tigre, Tigrinya, and Maltese.

The internal classification of the Afroasiatic phylum has long been a subject of controversy. Much of the evidence for genetic relationships within the phylum consists of comparative word lists from two or more families within the phylum. For an overview of such lists and their implications, see Frajzyngier and Shay (2012). Though some individual studies have been criticized for their methodology, they generally confirm the genetic unity of the Afroasiatic phylum (Frajzyngier and Shay 2012). In the present work, the internal structure of the phylum is taken as it is represented in Frajzyngier and Shay (2012).

Further evidence for the genetic unity of the Afroasiatic phylum comes from comparative studies of morphology. Most such studies have dealt with inflectional morphology, chiefly morphology of the verb (e.g. Diakonoff 1988). Though there is no uniform morphology across the phylum, some languages in all branches of the phylum exhibit inflectional and derivational morphemes with phonological structures and functions similar to those found in other branches. A number of these structures and functions are described later in this chapter. The only derivational morpheme whose reconstruction for Proto-Afroasiatic has been widely accepted is the prefix *m(V)-* (Greenberg 1963c, Lahaie 1985; see Sections 32.3.1 and 32.7.1).

One obstacle to postulating a common Afroasiatic morphology is that languages of all six Afroasiatic families have long been in contact with languages outside the phylum, including Niger-Congo, Nilo-Saharan, Sumerian, and Kurdish, with the result that some morphological features found in Afroasiatic may in fact be borrowed from other phyla or may reflect areal features (for a summary of language contact in Afroasiatic, see Frajzyngier and Shay 2012). There has also been considerable contact among languages of different families within the Afroasiatic phylum, with the result that features shared among two or more families may represent inter-familial borrowings rather than common retentions. Resolving such questions will require much more data on the undescribed and under-described languages of the phylum and considerable work on reconstruction of proto-languages within the phylum.

32.2 SCOPE OF THE CHAPTER

In descriptions of Afroasiatic languages, the term “derivation” may refer to the process of deriving one lexical item from another lexical item via affixation, suffixation, infixation, conversion, compounding, tone changes, vowel changes, and perhaps other means, or it may refer to the process of deriving a lexical stem or theme from an underlying abstract root. The latter process is known variously as root-and-pattern derivation (the term used in the present chapter), root-plus-vowel derivation, non-concatenative morphology, or templatic morphology.

The processes of root-and-pattern derivation and the forms and functions of the various themes resulting from this derivation have been well described for Semitic languages of all branches (Greenberg 1950, Hetzron 1997, Lipiński 2001, Holes 2004, Huehnergard

2011 (2005), and many others), Egyptian (Osing 1976, Reintges 1994, Loprieno 1995, and others), and Berber (Galand 2002, Vycichl 2005, Kossmann 2007, and others). The present chapter focuses on the means used in the six Afroasiatic language families to derive one lexical category from another lexical category (“class-changing derivation”) or to derive subcategories, for example, diminutive nouns, abstract nouns, or place names, from other lexical items within the same category. The Afroasiatic phylum is so large and so diverse that it is not possible to identify a common set of derivational means that characterizes the entire phylum. The goal instead is to provide an overview of derivational processes attested within the phylum. Some processes are shared by more than one family within the phylum, some appear (so far) to be unique to a given family, and some may be unique to a given language. There can be no complete description of derivational processes in Afroasiatic at this stage, given that many of the languages in the phylum have not been described.

The chapter is organized as follows: Section 32.2.1 provides a brief outline of root-and-pattern morphology as it has been described for Semitic, Egyptian, and Berber. Sections 32.3–32.6 provide an overview of derivational means in Afroasiatic. These sections are organized according to the lexical categories that are the outcome of derivation. At the editors’ suggestion, Section 32.7 is devoted to the description of derivational means attested in the Chadic language family, the most diverse family of the Afroasiatic phylum. Section 32.8 is a summary of the chapter.

32.2.1 Root-and-pattern Derivation in Afroasiatic

Root-and-pattern derivation, as described for Semitic, Egyptian, and Berber languages, begins with a root form, consisting of a discontinuous set of consonants and sometimes one or more vowels, which carries the basic semantic content common to all forms derived from the root. The pattern imposed on the root consists of a discontinuous set of one or more vowels, and sometimes a consonant. Consonants in the pattern applied to the root are often the result of gemination of one or more of the consonants in the root. In some cases the pattern may involve reduplication of the two or more consonants found in the root. The result of root-and-pattern derivation is a stem, or theme, that carries derivational and inflectional properties particular to a given lexical category. To the stem can be added other derivational means, including prefixation, suffixation, infixation, and compounding.

The most common form of the root in Semitic is a sequence of three obstruents, although roots with four obstruents are also attested, as are roots with only one or two obstruents (“weak” verbs; cf. Lipiński 2001). A frequently invoked example of root-and-pattern derivation is that of *ktb* ‘write,’ a root shared among Biblical Hebrew, Classical Arabic, and other Semitic languages. From this abstract root can be derived a large number of surface forms with related meanings, such as Classical Arabic *kita:b* ‘book’ (Noun), *kataba* ‘to write,’ *ʔaktaba* ‘to dictate,’ *taka:taba* ‘to write to each other,’ and many other forms (see Kaye 2007: 215–16), and Biblical Hebrew *kita:b* ‘book,’ *kotēb*

‘writing’ (participle), *ʔäktób* ‘I will write’ (imperfect tense) and other forms (Hetzron 1997: 63, 156). In Egyptian, the number of consonants or semiconsonants in the root varied from one to four. A vocalic pattern was superimposed on the root to form the stem, whose lexical category was a function of the vocalic pattern and the semantic content of the stem. Suffixes could then be added to the stem. Root-and-pattern derivation is also productive in the Berber family (Galand 2010 and others).

While the abstract lexical root in Egyptian is said to be of an indeterminate lexical category (Loprieno and Müller 2012), most studies of Berber and Semitic languages make a distinction between verbal roots and nominal roots. While it is theoretically possible for every verbal root to produce a noun via root-and-pattern derivation, there exist nouns in both Berber and Semitic that cannot be derived from verbal roots by any known pattern. Certain nouns in Semitic, sometimes called “primary nouns,” are taken to be derived from inherently nominal roots (Diakonoff and Kogan 2007). As stated in Gragg and Hoberman (2012), one piece of evidence for the existence of roots of different lexical categories is that all verbs, including loanwords, must conform to one of a limited set of canonical stem patterns, while nouns and adjectives, especially loanwords, do not have to conform to canonical patterns. In Berber, both nouns and verbs can be derived from most lexical roots. However, while all verbs have a nominal counterpart, not all nouns have a verbal counterpart, with the result that some roots can be considered inherently nominal (Galand 2010).

The remainder of this chapter is devoted to the means used in Afroasiatic languages to derive lexical categories from other lexical categories or to derive subcategories, such as diminutive forms, within a given lexical category.

32.3 DERIVING NOUNS

This section describes the various means attested in Afroasiatic for deriving nouns from other lexical categories. Although the term “noun” is not always defined in the source materials, it is taken to refer to a lexical item that can function, without additional marking, as the argument of a proposition.

32.3.1 Nouns from Verbs

All Afroasiatic languages described so far exhibit some means of deriving nouns from verbs. The evidence that a form is a deverbal noun is (a) that it is derived from a verb by some identifiable means, and (b) that it can be used as an argument in a predication. Among the categories of deverbal nouns found in languages across the phylum are place names, instrument nouns, agent nouns, nouns denoting actions or states, and other abstract nouns. The derivational means attested for deriving nouns from verbs include prefixation; suffixation, often accompanied by changes in vowel length, vowel quality, and tone; internal vowel changes; and reduplication.

All branches of the phylum show evidence of place names, instrument nouns, and agent nouns formed using the prefix *m(V)-*, the one derivational morpheme that is widely accepted as reconstructible for Proto-Afroasiatic (see Section 32.1). The prefix is productive in Chadic, Semitic, and Berber languages and is well attested in Egyptian and Coptic, but it is rare in languages of the Cushitic and Omotic families. In Semitic and Berber, it is not always clear whether the sequence *m(V)-* constitutes a prefix or is part of a pattern imposed on the stem (“preformant,” in Galand 2010). The sequence *m(V)-* does not appear to be productive in either Omotic or East Chadic, though there are a number of nouns that show evidence of a non-productive prefix. A number of examples are offered here:

Giziga (Central Chadic): *mùhúf* ‘farmer,’ from *húf* ‘to farm’; *mí-yí-dô-y* [mV-give birth-1SG.POSS-PL] ‘my parents,’ from *yí* ‘to give birth’ (Shay in progress).

Arabic (Semitic): *maktab* ‘office,’ *maktub* ‘printed material,’ *maktaba* ‘library’ from *ktb* ‘to write’; Akkadian *maškan-um* ‘location, place’ from *škn* ‘to place, to put.’

Tashelhiyt (Berber): *a-mzday* ‘inhabitant,’ from *zdy* ‘live in’; *a-mksa* ‘herdsman,’ from *ks* ‘graze’ (Kossmann 2012).

Egyptian: *mshn* ‘dwelling place,’ from *sni* ‘to inhabit’; *mḥt* ‘scales’ from *ḥ* ‘to measure’ (Lahaie 1985: 186).

Beja (also called Beḍauye; Cushitic): *mígnaf* ‘act of bowing, kneeling,’ from *genif* ‘to bow, to kneel’; *mahakīb* ‘act of sitting,’ from *hakeb* ‘to sit’ (Reinisch 1895).

Omotic: While there does not appear to be a productive prefix *m(V)-* in Omotic, there is some evidence of a non-productive prefix that may be associated with feminine gender, e.g. *miizza* ‘cow,’ *mac’cà* ‘female,’ *mácco* ‘wife’ (Azeb Amha p.c.).

While prefixes or preformants are the major means for deriving nominal stems from the abstract root in Semitic, Egyptian, and Berber, there are very few prefixes to be found in Cushitic, Omotic, and Chadic. Cushitic, which relies heavily on suffixes for deriving deverbal nouns, has the richest system of deverbal noun formation in the Afroasiatic phylum. Iraqw, spoken in Libya, has twenty nominalizing suffixes used to form nouns from verbs (Mous 1993: 74–82), although not all suffixes can be used with all verbs and not all suffixes have distinctive meanings. For K’abeena, spoken in Ethiopia, Crass (2005) describes twelve distinct and productive deverbal nominalizers and a few others that are less productive. The Sidamo language of southern Ethiopia has almost as many deverbal nominalizers (Anbessa 2000: 67–9).

Mous (2012) cites the agentive noun, formed via suffixation, as common across Cushitic languages. Examples are Sidaama *hat’ar-aančo* ‘butcher,’ from *hat’ar* ‘to butcher’ (Anbessa 2000: 67), and the productive agentive suffixes *-usmo* (male agent),

-*uso* (female agent) and -*use* (plural agent) in Iraqw (Mous 1993). Other functions of nominalizing suffixes in Cushitic include deriving instrument nouns and abstract nouns from verbs, for example, Iraqw *waawiti'iim-a* 'rule' (abstract feminine noun) from *waawiti'iim* 'to rule' (for similar examples in Sidaama, see Anbessa 2000: 7, cited in Mous 2012). In some languages there is overlap between deverbal and denominal nominalizers (see 32.3.1.2). In some Cushitic languages nouns may be derived from verbs via internal vowel changes alone: Beja *ba:r* 'act of snatching' from *bir* 'to snatch'; *da:f* 'act of going' from *dif*; Tuareg *e-fi:l* 'act of leaving' from *efel* 'to leave' (Vycichl 1988).

The prefix or preformant *a-* is productive in Touareg, where it derives a noun from the causative verb form, itself marked by the preformant *s-*: *asədəg* 'act of stinging,' from *sədəg* 'cause to sting' (cf. *ədəg* 'to sting') (Galand 2010: 150).

A few nominalizing suffixes are also found in Chadic languages (see Section 32.7.2).

32.3.2 Nouns from Nouns

While a number of nouns in Afroasiatic languages show evidence of derivation from other nouns, this path of derivation is not highly productive. The coding means attested include affixation and compounding.

In Cushitic, some of the same suffixes used to derive nouns from verbs can also derive nouns from other nouns, for example, *daqari'iima* 'prosperity' from *daqarmo* 'rich person' (Mous 1993: 77; see Section 32.3.1); cf. *but'-ima* 'poverty' from *but* 'to be poor.' Semitic languages have several means for deriving abstract nouns from concrete nouns, including infixation and suffixation, for example, Akkadian *awīl-ūt-um* 'humanity, mankind' (from *awīl-um* 'man'), Israeli Hebrew *ʔenošut* 'humanity' (from *ʔenós* 'human being' and *xovevanút* 'amateurism' (from *xoveván* 'amateur') (Gragg and Hoberman 2012).

Compounding, a common cross-linguistic means of deriving new nouns from existing nouns, is uncommon in Afroasiatic. Berber languages have a few compound nouns for plant names and body parts, for example Eastern Riffian *a-ɣəsmir* 'lower jaw' ('beard-bone'; cf. *iyəss* 'bone' and *θ-mar-θ* 'beard'); *a-ɣəzdis* 'rib' ('belly-bone'; cf. *iyəss* 'bone' and *a-ʕəddis* 'belly') (Kossmann 2012).

In Earlier Egyptian (3000–1300 BCE), nouns were derived from the verbal stem by suffixation and the addition of nominal morphology. In the transition from Earlier Egyptian to Later Egyptian, nouns derived from the stem were replaced by lexicalized compounds (Till 1970, as cited in Loprieno and Müller 2012), possibly as a result of phonological changes.

For the derivation of nouns in Chadic languages, see Section 32.7.

32.3.2.1 Diminutives

Languages of all Afroasiatic families have means of deriving diminutive nouns from other nouns. In Semitic languages, diminutives can be derived from other nouns by imposing

a different vowel pattern on the root or by suffixation. Diminutive suffix forms include Tigre *-ay* (*bab* ‘door,’ dim. *beb-ay*; *bet* ‘house,’ dim. *bet-ay*); Mehri *-ōt* (*rəhbēt* ‘place,’ dim. *rəhbānōt*; *rəhmēt* ‘vegetation,’ dim. *rəhmānōt*); and Aramaic and Hebrew *-o:n*.

In Zenaga (Berber), diminutives of some masculine nouns are formed via circumfixation, using the prefix *ay-* and the suffix *-t*. This is sometimes accompanied by internal vowel changes. The combination of means suggests root-and-pattern derivation rather than affixation: *ay-əgrar-t* (son-ram-SG:F) ‘small ram’ (cf. *əgrar* ‘ram (m.),’ *ay-wa’r-t* ‘small lion’ (cf. *wa’r* lion) (Taine-Cheikh 2002). A few diminutives are formed through reduplication of the root, e.g. *tefēlfēlt*, diminutive of *tefilt* ‘flat rock.’

Gender marking is used in Chadic, Cushitic, and Berber languages to indicate small size, and in some cases pejorative meaning. The masculine gender typically refers to larger entities than the feminine gender, for example Figuig (Berber) *a-yanža* (m.) ‘ladle,’ *ta-yanžay-t* (f.) ‘spoon’ (Kossmann 2012). Most Omotic languages show the same pattern, though the Hamar language appears to be an exception. According to Lydall (1988: 78), feminine gender in Hamar “is used to indicate large and major things... while masculine gender is used to indicate small and minor things.” A similar phenomenon is found in Puvé (Chadic; Shay, submitted), where the phrase “mother”+Noun codes importance or large size, as in *gu* ‘tree’ vs. *ya gu* ‘large/important tree’ (lit. ‘mother tree’). Importance or size can also be coded by use of the feminine agentive prefix *i-*, as in *gway?* ‘mountain’ vs. *i-gway?* ‘large/important mountain.’

In Migama (Chadic; Junggraithmayr and Adams 1992), a change in gender in either direction codes the diminutive: The diminutive variant of a masculine noun is marked feminine, and the diminutive variant of a feminine noun is marked masculine:

(1) <i>gáàpú</i> (m.) ‘large horn’	<i>gáàpè</i> (f.) ‘small horn’
<i>dâmbá</i> (f.) ‘mountain, mountain range’	<i>dâmbú</i> (m.) ‘rock’

32.3.3 Nouns from Other Sources

Derivation of nouns from adjectives is widely attested in Cushitic languages but is not common in other Afroasiatic languages, though it is attested in Wandala (Central Chadic; Frajzyngier 2012). In Sidaama (Cushitic), the same suffix that derives abstract nouns from verbs and from other nouns (Sections 32.3.1 and 32.3.2) may also derive nouns from adjectives, for example *danč-imma* ‘goodness’ from *danča* ‘good’ (Anbessa 2000: 69, as cited in Mous 2012). In Iraqw, the related suffix *-eema* is productive in deriving nouns from adjectives: *saw* ‘far,’ *saweema* ‘distance,’ *hhoohhoo* ‘beautiful,’ *hhoohhooèema* ‘beauty’ (Mous 1993: 82).

K’abeena (Cushitic) has a nominalizing suffix *-iti* that derives nouns from ideophones (Crass 2005: 84, 233).

32.4 DERIVING VERBS

Afroasiatic languages exhibit only a few productive means for deriving verbs from other lexical categories. These means are described in the following sections. The term “verb” is taken to mean a lexical item that can serve, without additional marking, as the predicate of a proposition.

32.4.1 Verbs from Nouns

While all Afroasiatic languages have means for deriving nouns from verbs, the reverse process—deriving verbs from nouns—is far less common. In Cushitic languages, some of the same suffixes that derive nouns from verbs can also derive verbs from nouns. According to Mous (2012), the inchoative verb, used to express the notion of entering into a state, entering into a social position, or acquiring a certain property, can be morphologically derived from either a noun or adjective by means of the same affix. Examples from Wolaitta, where the inchoative is marked by *-t-*: *kawo-t-iīsi* ‘he became a king,’ from *kawó* ‘king’; *gita-t-iīsi* ‘he/it became big,’ from *gitá* ‘big’; and *dure-t-iīsi* ‘he became rich,’ from *dúre* ‘rich’ (Mous 2012). In Somali, the factitive suffix *-ayn* or *-ays* also derives verbs from nouns (Saeed 2007: 568). In Iraqw, the inchoative suffix *-uw* occasionally derives a verb from a noun, for example, *xwayla* ‘birth,’ *xwayluw* ‘to give birth’ (Mous 1993: 187). The suffix is more commonly used to derive verbs from adjectives (*tleer* ‘long,’ *tleeruw* ‘to become long’) (Mous 1993: 187).

In Ouargla (Berber) the factitive prefix *ss-*, apparently cognate with the causative marker in other Afroasiatic languages (cf. Greenberg 1963c), may derive a verb from a noun, for example *ss-iwəl* ‘speak’ from *awal* ‘word’ (Kossmann 2007).

Egyptian shows evidence of so-called root-to-root derivation, where a verb can be derived from a nominal or verbal root through reduplication (cf. Reintges 1994):

(2) Root	Reduplication
<i>sn</i> ‘brother’	<i>snsn</i> ‘to befriend’
<i>gmj</i> ‘to find’	<i>ngmgm</i> ‘to be gathered’ (with reflexive <i>n</i> -prefix)
<i>snb</i> ‘to be healthy’	<i>snbb</i> ‘to greet’

There is some evidence that in Hebrew and Arabic, and perhaps in other Semitic languages, the consonantal root may be extracted from a borrowed noun and a verbal pattern may be imposed on the noun to form a new verb, as in *telefon* ‘telephone’ (N), *tilfen* ‘(telephoned)’ (Hebrew; Goldenberg 1994, see also Chapter 12).

32.4.2 Verbs from Ideophones

Ideophones, characterized by phonological forms unusual for the language in question and by constrained distribution within the language, are a common category in languages throughout Africa, both within and outside of the Afroasiatic phylum. Ideophones are well attested in the Chadic, Cushitic, and Omotic families. Derivation of verbs from ideophones is very common in the languages of southern Ethiopia. In a productive derivational process that has been identified as an areal feature of southern Ethiopia, a verb corresponding to “to say” (or, less commonly, the verb “to do” or “to put”) may be combined with an ideophone to produce a compound verb (see Leslau (1945) and later works and Hetzron (1972)). In the Omotic languages, the verb corresponding to “say” plus an ideophone produces an intransitive predicate (Azeb Amha 2012) while the verb “to do” plus an ideophone produces a transitive predicate: Wolaitta *páški ga* ‘be happy’ (*páški-g-a* IDEO-say-IMPER); *páški ?oottá* ‘make happy!’ (IDEO do-IMPER). The same process is productive in Amharic (Ethiopian Semitic): *šwa alä* ‘fizz’ (*alä* ‘to say’), *futt alä* ‘slurp’, *həqq alä* ‘hiccup’, *suləll alä* ‘run here and there’, *qučč alä* ‘sit down’, *zəmm alä* ‘be quiet’ (Gragg and Hoberman 2012).

Appleyard (2001) and others posit Cushitic languages as the source of the ideophone + “to say” construction, which Appleyard terms “composite verbs.” Some examples of composite verbs in Cushitic are: *čal jedē* ‘be quiet’, *lip jedē* ‘to blink’ (Oromo); *hawwu yi* ‘feel dizzy’, *sammi yi* ‘be quiet’ (Sidamo); *har yiri* ‘stand up’, *tuš yiri* ‘to prick’ (Somali) (all from Appleyard 2001). South Cushitic also has a suffix *-eel* that derives verbs from ideophones (Mous 2012).

Ideophones are common in Chadic languages, often serving as nominal or verbal modifiers. In the Chadic language East Dangla, an ideophone can function as a verb through reduplication alone, for example *ɲà súyúp-súyúp-súyúp* (3M IDEO-IDEO-IDEO) ‘to run away’, where *súyúp* represents the sound made by the runner. This form co-exists with the verbal phrase *ɲà gadā* ‘he ran away’, from *gādē* ‘to run’ (Shay 1999; see also Section 32.7.3).

32.5 DERIVING ADJECTIVES

Whether or not adjectives constitute a distinct lexical category is highly debatable for many Afroasiatic languages. According to Kossmann (2012), Tashelhiyt has only one adjective, *lždid* ‘new’ (a borrowing from Arabic), while Tuareg lacks the category adjective altogether, coding nominal modification via relative clauses using stative verb forms. Northern and Eastern Berber languages have a set of nominal forms that can function either as head nouns or as modifiers, suggesting that adjectives are a subcategory of nouns.

A similar dilemma exists for Cushitic languages. According to Mous (2012), adjectives in Ts’amakko belong to the class of nouns, while adjectives in Afar are stative verbs, a claim that has also been made for Berber adjectives (Taine-Cheikh 2003). In Somali (Saeed 2007), a small set of inherent adjectives is enhanced by adjectives derived from verbs using the stative suffix *-an*, similar to the means used in Tuareg. In the Omotic

family, adjectives share morphosyntactic properties with both nouns and verbs (Azeb Amha 2012). In some Chadic languages, such as Mupun and Lele, adjectives form a subcategory of verbs (Frajzyngier 1993), with the modifying function being coded by the use of the stative verb form in a relative clause.

The following sections describe only adjectives that have clearly been derived from another lexical category, where “adjective” is defined as a lexical item whose unmarked function is to modify a noun. Such adjectives may belong to a separate lexical category, or they may belong to a subcategory of nouns or verbs coding property concepts.

32.5.1 Adjectives from Nouns

In many Afroasiatic languages, adjectives are derived from nouns through productive or non-productive suffixation. Ts’amakko (Cushitic) has a suffix *-(V)l* that derives adjectives from nouns, as in *?us-kakk-ol-akka* ‘dirty’ from *?us-kakk-o* ‘dirt’ (Savà 2005: 93). In Egyptian, a morpheme **ij* (masc.) or **it* (fem.) was added to a noun stem to form a corresponding adjective, in a process called nisbation: *ntr* **/na:caɪ/* ‘god’, *ntrj* **/nuc.ij/*, *ntrj.t* **/nuc.it/* ‘divine’ (Loprieno and Müller 2012).

Semitic languages have a highly productive suffix for deriving adjectives, often called relational adjectives, from nouns. The suffix appears in Arabic as *-i:y*, as in *šams* ‘sun’, *šamsi:y* ‘solar, photo-’; *madrasa* ‘school’, *madrasi:y* ‘scholastic’ (Holes 2004: 160); in Akkadian as *-i*, as in *šap-lum* ‘bottom’ *šaplī-um* > *šaplûm*, *šaplī-tum* ‘lower (m,f)’; in Hebrew as *-i* and in Syriac as *-a:y* (Gragg and Hoberman 2012) (see also Section 32.7.2).

32.5.2 Adjectives from Verbs

In Ts’amakko (Cushitic), even though adjectives form a subcategory of nouns, some adjectives show evidence of derivation from verbs. Savà (2005: 93) identifies eleven adjectives derived from verbs via the suffix *-ay*. This suffix is followed by another suffix coding gender or number, evidence that the adjectives belong to the noun class despite their verbal origin: *daaf-akko* ‘blind’ (m.), *daaf-atte* ‘blind’ (f.), *daaf-ayke* ‘blind’ (pl.), from *daaf* ‘to be blind.’ South Cushitic has a non-productive suffix *-ar* that appears in adjectives derived from verbs. In Egyptian, semivocalic suffixes derive adjectival forms (participles and relative forms) from the verb stem (Loprieno and Müller 2012).

32.6 DERIVING ADVERBS

The status of adverbs, like that of adjectives, as a distinct lexical category is not obvious in all Afroasiatic languages. Mous (2012) states that adverbs are not a clearly defined

word-class in Cushitic. Where a class of adverbs can be identified, adverbs across the Cushitic family show evidence of derivation through suffixation, for example Alagwa -*nkoo*: *bankóo* ‘first’, *hinkóo* ‘now’, *lankóo* ‘some time, some day’, *ṭankóo* ‘before’, and -*ee*, *baree* ‘if’, *tsigée* ‘fast, early’, *tsobolee* ‘truly’ *sigee* ‘far’ (Mous 2012).

In Ouargla (Berber), adverbs can be derived from nouns via the suffix -*i*, often accompanied by omission of the nominal prefix *a-*: *a-rgaz* ‘man’, *rgaz-i* ‘man-like’; *kukām* ‘silence’, *kukm-i* ‘silently’ (Kossmann 2012).

In Giziga (Chadic), the adverb “alone” is derived from the numeral *pál* ‘one’ followed by a possessive suffix coding the person and number of the subject: *pál-dū* ‘I alone’ (lit. ‘my one’; Shay, in progress).

32.7 DERIVATION IN CHADIC LANGUAGES

The following sections describe derivational means attested in Chadic languages and the functions they encode. It is not possible to produce a complete survey of word formation in Chadic, as only about one-fourth of the Chadic languages have been described, and many of these descriptions are only brief sketches. The goal of this section is to provide an overview of derivation in Chadic using examples from all branches of the Chadic family: West, Central (also called Biu-Mandara), East, and Masa (sometimes considered a sub-branch of Central/Biu-Mandara).

In some Chadic languages, major lexical classes show evidence of derivation via means that are no longer productive. For example, in many Chadic languages, all verbs begin with a consonant, while nouns and other lexical categories can begin with either a vowel or a consonant. In some Chadic languages the underlying form of the verb consists solely of consonants while the underlying form of the noun consists of both consonants and vowels. In other languages, verbs have a limited set of vowels while nouns have a much broader set. While the focus of the present chapter is on derivational means that are productive and/or transparent, evidence of earlier derivational means that are no longer productive is mentioned as well. When possible, examples of each derivational means are given for each branch of the Chadic family.

32.7.1 Prefixation

The prefix *m(V-)* (see Sections 32.1 and 32.3.1) is attested across the Chadic language family. The prefix is productive in the West and Central branches. The prefix is also evident in many nouns in East Chadic, though it is not obvious that the prefix is productive. Typically, the productive prefix derives an agent, instrument or locative noun from a verbal stem. In Hausa (West Chadic), for example, the prefix *m(V-)* may derive an agent noun from a verb: *hàifā* ‘give birth to’, *mahàifā* ‘parent’ (m.), *mahàifiyā* ‘parent’ (f.); *bī* ‘to follow’, *mabiyā* ‘follower (m.)’, *mabiyiyā* ‘follower (f.)’ (Jaggar 2001: 108). The deverbal noun prefix has been

identified in other West Chadic languages, including Margi (Hoffmann 1963), Bade (Schuh 2007), and Goemai (Hellwig 2011).

The prefix *m(V)*- is widespread in Central Chadic languages. As shown in Section 32.3.1, the prefix is productive in forming nouns from verbs in Giziga (Central). The same prefix also derives nouns from other nouns: *màdàngwàs* ‘old woman’ (*dà* (term of respect for a woman) + *ngwàs* ‘woman’), *màzínhàn* ‘fingers’ (*hàn* ‘hand’), *màlàk* ‘stranger’ (*làk* ‘visit (N)’), and adjectives from nouns: *móngósò* ‘female’ (*ngwàs* ‘woman’), *màzlè* ‘male’ (*zàl* ‘man’). Many Giziga nouns whose path of derivation is no longer obvious also begin with *m(V)*-. *màzàwàl* ‘castrated goat’, *mòtòkònyò* ‘hyena’, *mòzòròb* ‘anus’, *migisin* ‘ant’, *méngévèl* ‘wasp’, *mikléf* ‘boule’, *mándàžày* ‘rocks’, *màkàkwá* ‘border’, *mèzègwèl* ‘custom, ritual’ (Shay, in progress, Frajzyngier and Shay, field notes). Very similar patterns are found in Gidar (Central), where the prefix *m(V)*- derives agent nouns from verbal stems, for example *hál* ‘steal’, *mà-hálá* ‘thief’, and is also common in attributive nouns, for example *má-zànlákàn* ‘male’, *mú-gùlùngú-nì* ‘adult female animal’ (Frajzyngier 2008). Like Giziga, Gidar has nouns that may have a prefix *m(V)*- but whose derivation is not obvious: *móngòsòk* ‘broken piece of calabash used as a ladle’, *méptèkè* ‘plain mush, without gravy’ (Frajzyngier 2008). In Mofu-Gudur (Central), the prefix *m(V)*- derives the verbal noun or infinitive from the verb stem: *mézemey* ‘meal’, ‘act of eating’ (*zəm* ‘to eat’). It also derives ordinal numbers from cardinal numbers, for example *mécew* ‘second’ from *cew* ‘two’, and derives a large variety of nominal forms from other verbal and adjective-adverbial stems (Barreteau 1988: 161). The prefix *m(V)*- in Wandala (Central) is found in nouns designating places, languages, and people: *múfàkè* ‘Kanuri language’, *màsɣáyá* ‘Giziga language’, *múyánè* (village inhabited by speakers of Podoko) (Frajzyngier 2012).

In Masa (Central or Masa branch), agent nouns can be derived from the verbal root by the prefix *m(V)*- and a suffix *-na* (m.) or *-ta* (f.), which functions as a determiner (“spécificateur,” in Melis 1999) when attached to a noun. The agent noun is often preceded by the noun *sá* ‘man, person’: *sà má sùn-tà* ‘messenger’ (*sun* ‘to send’), *sà mà dūr-rà* ‘warrior’ (*dūr* ‘to fight’).¹ In East Chadic, the large number of agent and instrumental nouns with initial *m(V)*- suggests that the prefix was at one time productive.

A few prefixes other than *m(V)*- are attested in Chadic. Hausa has a productive prefix *mai*- ‘owner of’, ‘one who’, that derives nouns from verbs and other nouns: *maisāyār* ‘seller’ (‘one who sells’); *maitāfiyà* ‘traveller’ (‘one who travels’); *maidākī* ‘wife’ (*dākī* ‘room, hut’); *maigidā* ‘head of the household’ (Newman 2007b: 145). Hausa also has a prefix *ba*- that derives ethnonyms, that is, nouns referring to ethnicity, origin, occupation, or other attributes, from toponyms or other nouns (Jaggar 2001: 113ff.): *bàfādè* (m.), *bàfādiyā* (f.) ‘courtier’ from *fādā* ‘palace’, *Bàsakkwacè* (m.), *Bàsakkwaciyā* (f.) ‘Sokoto person’, from *Sakkwato* ‘Sokoto area (Nigeria)’. The prefix *àn*- in Pero (West), like *m(V)*- in other Chadic languages, derives agent nouns from verbs and other lexical categories (Frajzyngier 1989): *àn-dīmmà* ‘beggar’ from *dīmmò* ‘to beg’, *àn-lè* ‘parent’ from *lé* ‘to give

¹ Not all of the symbols used in Melis (1999) could be replicated in the present chapter. The importance of the examples does not lie in the phonological forms but in the fact that the various derivational means are attested in the Masa branch.

birth.' The same prefix derives ordinal numerals from cardinal numerals, for example, *àn-bèlòw-i* 'second' from *bèlòw* 'two' (the suffix *-i* marks the noun as definite). Bade and Miya (West) have a similar prefix *a-* that derives nouns from verbs: *àlhu* 'say', *alhán* 'speech'; *mànu* 'spend a night', *àmanən* 'year' (Bade, Schuh 2007); *átám* 'song', *təma* 'sing'; *ákír* 'theft', *kəra* 'steal'; *ázhigə* 'praise' (N), *zhiga* 'to praise' (Miya, Schuh 1998). In Wandala (Central), a prefix *lə* 'one of' derives human nouns from other nouns: *lə ksè* 'chief, sultan' (lit. 'one of town'); *lə hálé* 'mean person' ('one of bad mood') (Frajzyngier 2012).

A few prefixes are also found in East Chadic. In Mokilko the prefix *ta-* may derive an abstract noun from concrete noun, as in *t-báraŋa* 'kingdom' from *báraŋa* 'king.' Some color terms in Somrai are characterized by the prefix *gV-*, for example *gezáge* 'blue', *gusé* 'green', *guámgo* 'red', but it is not clear whether the prefix is productive. A similar prefix is found in the Kwang dialect Modgel: *kasaragán* 'blue', *kabasán* 'green', *kéletən* 'red' (all data from Lukas 1937b).

In a number of West Chadic languages, animate nouns and nouns referring to body parts appear to have a non-productive nasal prefix, probably a borrowing from non-Chadic languages spoken in the same area. The nasal prefix is usually homorganic with the following consonant: Mupun *mvér* 'lion', *njar* 'termite', *ŋám* 'ram'; Gwandara *ñcucuwa* 'bird', *ŋkuciya* 'dove' (Frajzyngier and Koops 1989: 236–7). A similar prefix is found in Sura (Mwaghavul; Jungraithmayr 1963) as well as in Hausa.

32.7.2 Suffixation

Suffixation is more prevalent in Chadic than prefixation. Chadic languages of all branches have productive suffixes for deriving nouns from verbs, adjectives, and other nouns. As noted in Section 32.3.1, derivation of nouns from verbs is attested in all Chadic languages described so far. Many such derivations are products of suffixation.

In Pero (West), the suffix *-áani* can be added to a verb stem to form a noun (*nomen actionis*) that may function as an argument, complement, or stative predicate: *káp* 'to tell', *kápáani* 'telling (N)'; *káur-* 'to feed', *kàuráani* 'feeding'; *yú* 'to die', *yùwáani* 'death' (Frajzyngier 1989: 125–126). Some nouns have the *-áani* suffix when used in the future tense and a different nominalizing suffix, either *-i* or *-ò*, when used in the nominal function in other constructions.

In Miya (West), some verbs have as many as three verbonominal counterparts, labeled in Schuh (1998) as participle, gerund, and deverbal noun. The participle, which occurs only as part of a larger verbal construction, is marked by the suffix *-a(w)*. The gerund, which can function as an argument, is marked by the suffix *-ahiy* or *-akə*. The forms Schuh labels deverbal nouns are marked by prefixation (see Section 32.4.1). Hausa (West) also has a rich system of derivational suffixes (see Newman 2007a and others).

In Hdi (Central), verbal nouns are formed via the suffix *-u* or *-i*: *ghálá* 'steal', *ghálú* 'theft'; *và* 'light a fire', *vú* 'fire (N)'; *xíídá* 'to bite', *xíídí* 'bite (N)' (Frajzyngier with Shay 2002). In Wandala (Central), some abstract nouns are derived from adjectives via the morpheme *-irè* 'head' (Frajzyngier 2012): *tàpdà mmè* 'dark', *tàpdà mm-irè* 'darkness';

pàràké 'light,' *pàràkí-irè* 'lightness'; *jírè* 'true,' *jír-irè* 'truth.' In Masa, the determiner *-na* (m.) or *-ta* (f.) suffixed to the verbal root derives abstract nouns from verbs: *mà* 'to come,' *mà-nà* 'arrival'; *zàl* 'to heat,' *zàl-là* 'heat (N)'; *mit* 'die,' *mit-tà* 'death' (Melis 1999: 120ff).

East Dangla (East, Shay 1999) has a productive suffix *-iya* that forms abstract nouns from verbs: *bùgúmé* 'to be quiet,' *bùgúmiyà* 'silence'; *sòone* 'to dream,' *sòoniya* 'dream (N)'; and a suffix *-aw* that derives nouns from verbs and other parts of speech: *tátá* 'large,' *tataw* 'largeness,' *bako* 'neighbor,' *bákaw* 'neighborhood,' *gàdè* 'to run,' *gadaw* 'run (N)'. In Bidiya (East, Jungraithmayr 1990), an abstract noun and an agent noun can be derived from the same verb stem using different suffixes: *zawàl* 'to hunt,' *zawàlò* 'hunt' (N), *zawàalà* 'hunter'; *riy* 'to work,' *riyò* 'work (N),' *riyàawà* 'worker.' Mokilko (East, Jungraithmayr 1990) has a similar suffix *-o*, e.g. *nàab-ó* 'work (N)' from *nàabé* 'to work,' as well as a suffix *-kó/-ká*, which derives language names and abstract nouns from concrete nouns and adjectives. In Migaama (East, Jungraithmayr 1992), the suffix *-i* derives abstract nouns from adjectives and concrete nouns: *áyáawí* 'wealth' from *áyáawí* 'rich'; *dábàali* 'work (N)' from *dábàali* 'worker (f.)'. In Bidiya (East, Alio and Jungraithmayr 1989) the suffix *-ga* derives color adjectives from verbs: *sópòr* 'to become white,' *sópòrga* 'white'; *baar* 'to become red,' *barga* 'red.' Mokilko has a similar suffix *-kó/-ká* that forms abstract nouns from adjectives and from nouns denoting personal or other characteristics: *sò'òró* 'long,' *sò'ir-kó* 'length'; *lánzé* 'friend,' *lánzi-ká* 'friendship'; *mèedá* 'large,' *mèedí-ká* 'largeness,' *tùwàrwá* 'white,' *tùwàr-ká* 'whiteness.' The same suffix also derives names of languages from toponyms: *móòkil-kó* 'language of Mukulu' (all examples from Jungraithmayr 1990).

The categories adjective and adverb in many Chadic languages often show evidence of derivational morphology via suffixes that are no longer productive. In Mupun (West; Frajzyngier 1993), a large proportion of the small set of adverbs end in *-k*: *rák* 'only,' *dúk* 'also,' *kùk* 'only, just.' A number of property-concept terms in Mina (Central) appear to be derived through the suffix *-ék*: *kwèdék* 'white,' *gùzandék* 'red,' *dèlék* 'green, blue' (Frajzyngier et al. 2005). Giziga (Central) has a number of property-concept terms apparently derived through the suffix *-aŋ*: *gàgàzàŋ* 'red,' *tàtràŋ* 'black,' *mádà dāŋ* 'large' (Shay, in progress). Bidiya (East) has a similar suffix used in some deverbal nouns: *soy* 'to say, shout,' *soy-èn* 'saying, shouting'; *waar* 'to dance,' *wáar-eŋ* 'dance (noun)' (Alio and Jungraithmayr 1989). In Wandala (Central), most adjectives end in the vowel *-e*, possibly related to the final *-i* found in many adjectives in Semitic (see Section 32.5.1): *nàmlàk* 'dirty,' *dónwè* 'black,' *zùŋwè* 'first' (Frajzyngier 2012).

A suffix *-kə*, with polar tone, marks the diminutive or the feminine gender in Gidar (Central): *gámdá* 'chicken,' *gámdá-k* 'small chicken'; *sàrá* 'white man,' *sàr-kə* 'white woman' (Frajzyngier 2008).

32.7.2.1 Vowel and Tone Changes

As indicated by many of the examples in Section 32.7.2, a derivational suffix is very often accompanied by stem-internal vowel and/or tone changes (see Frajzyngier 1981). In Lele (East), where verbal nouns are derived from the past tense form of the verb (Frajzyngier 2001), the noun is derived from the verb by lowering the final vowel, provided the past

tense form ends in *-i*. The tone on the lowered vowel is determined by the syllabic and tonal structure of the verbal root: *òjì* ‘help’ (past tense), *òjè* ‘help’ (N); *dènglì* ‘to hear,’ *dènglé* ‘hearing.’ If the past tense form ends in a consonant, the vowel *-e* is added to derive the verbal noun: *kàl* ‘climb’ (V), *kale* ‘climb’ (Noun); *kin* ‘return (Past),’ *kinè* ‘return (N)’ (Frajzygier 2001: 43ff). Similar noun–verb pairs involving both suffixed vowels and stem-internal vowel and tone changes are found in Giziga (Central), for example *jíni* ‘to work together,’ *jiné* ‘common work (N)’ (Shay, submitted), and throughout the East branch: *dàbállò* ‘to deceive,’ *dábáalí* ‘deception’ (Migama, Jungraithmayr and Adams 1992); East Dangla *óoyé* ‘to greet,’ *òoyá* ‘greeting (N)’; *gàwne* ‘to cultivate’ *gawni* ‘act of cultivating,’ *dyàawe* ‘to walk,’ *dyiwi* ‘walk (N)’ (Shay 1999); Bidiya *eyem* ‘to eat,’ *eyémò* ‘act of eating’ (Alio and Jungraithmayr 1989).

In a few instances a deverbal noun is formed by means of internal vowel and/or tone changes without suffixation: East Dangla *gidíyé* ‘to trade,’ *gidây* ‘purchase (N)’; Miya *ràvadə* ‘thinness,’ *rəvadə* ‘to grow thin’; *tsàwədə* ‘carving,’ *tsooda* ‘to carve’ (Schuh 1998: 117).

32.7.3 Reduplication as a Derivational Means

There is ample evidence of productive and non-productive reduplication throughout the Chadic family. Reduplication as a means for deriving adjectives and adverbs from other categories is found in all branches. In some languages reduplication is also a means of deriving nouns from verbs. Reduplication is common in the formation of ideophones (see later in this section), and there is some evidence of non-productive reduplication in some verbal and nominal stems. Both complete and partial reduplication are attested in Chadic. In partial reduplication, it is often the leftmost two or three segments that are reduplicated. There are exceptions, such as reduplication of the rightmost syllable in East Dangla (East), along with vowel lengthening, to produce a diminutive form: *iyà* ‘Mother’ (form of address), *iyàayá* ‘Mommy’ (form of address); *kòkìra* ‘chickens,’ *kòkìráarà* ‘baby chicks’ (Shay, field notes).

Reduplication is productive in several domains in Mupun (West, Frajzyngier 1993). Verbal nouns are derived via complete reduplication of the verb: *kám* ‘to teach,’ *kàmkám* ‘teaching’; *rān* ‘to write,’ *rānrān* ‘writing.’ Some agent nouns are also formed via reduplication: *mān* ‘to know,’ *mānmān* ‘knowledgeable person.’ Many adverbs are derived from adjectives or nouns via reduplication: *bál* ‘strong,’ *balbal* ‘strongly,’ *sù* ‘to run,’ *səsù* ‘quickly,’ *kénkén* ‘quickly’ (identical with the plural and intensive forms of the adjective). In Miya (West), some deverbal nouns are derived by means of reduplication: *tyákyáki* ‘scratching’ (*tlakə* ‘to scratch’), *didyadi* ‘falling’ (*dəda* ‘to fall’) (Schuh 1998). Miya also has adjectives derived via reduplication, for example, *gózí-gózíny* ‘weak.’

In Hdi (Central), some color terms are derived via reduplication of noun forms (Frajzyngier with Shay 2002): *kùzún-kùzún* ‘green’ (*kùzún* ‘fresh leaves’), *rbihl-rbihl* ‘gray’ (*rbihl* ‘clay’). Whole or partial reduplication of the noun also characterizes some color terms in Giziga (Central), where reduplication may be accompanied by use of the

suffix *-àŋ*: *kàlàf-kàlàf* ‘green’ (*kàlàf* ‘leaf’); *bàbàràŋ* ‘white’ (*bàràŋ* ‘white man’); *gàgàzàŋ* ‘red, brown’ (*gàzàŋ*, also ‘white man’), *zàyzùŋ-zàyzùŋ* ‘yellow’ (*zà* ‘feces,’ *zùŋ* ‘child’).

Adverbs in all branches of Chadic also show evidence of reduplication, though the source is not always evident, for example, Giziga *pátpát* ‘many times,’ *kédkéđ* ‘for a long time’; Hdi *misi-misim* ‘fast, easily,’ *tántán* ‘first,’ *xótxótà* ‘certainly,’ *ndándáná* ‘immediately.’ When the Hdi adverb is derived from a verb, the verb is reduplicated and each reduplicated segment is preceded by the locative preposition *tà* (Frajzyngier with Shay 2002: 234):

- (3) *mǝhám tá ghùn-àf-t-íí tà xwáyá tà xwáyà*
 chief COM send-UP-REF-1SG PREP run PREP run
 ‘it is the chief who sent me urgently’

Formation of an adverbial via reduplication of both verb and preposition is also attested in Mupun (West):

- (4) *mo se lua əh si-ɖaar si-ɖaar*
 3PL eat meat goat PREP-day PREP-day
 ‘They eat goat meat daily’ (Frajzyngier 1993: 275)

In Mina (Central), the adverb *cíké cíké* ‘together’ is derived from the quantifier *cíké* ‘all’ via reduplication, the adverb *bibíc-bibíc* ‘all day’ is derived from the noun *bibíc* ‘day,’ and the adverb (*n*)*tátà* ‘alone, only’ is derived via reduplication from the numeral *ntá* ‘one’ (Frajzyngier et al. 2008). Double reduplication of the noun *cidě* ‘pile’ derives the adverb of manner ‘in piles’:

- (5) *wirnjik dīy-à bək-áhà*
 ash put-GO pour-GO
cidě cidě cidě cidě á kàtəf
 pile pile pile pile PRED road
 ‘Ash was pouring out of the shoe in small piles on the road.’ (Frajzyngier et al. 2005: 166)

In Lele (East), the adverb *samsam* ‘slowly’ is derived via reduplication from the quantifier *sam* (Frajzyngier 2001), and the adverb *téŋtéŋ* ‘on his own, separately’ is derived from the quantifier *téŋ* ‘alone.’ The East Dangla (East) adverb *kàk-kàràk* ‘separately’ is apparently derived via reduplication, though the source is not evident.

Partial reduplication of a modifier may code intensification of the property lexicalized by the modifier, as in East Dangla *kódǝki* ‘quickly,’ *kók-kódǝki*; *ànáŋ* much, *ànáŋ* ‘very much.’ In Hausa, in contrast, complete reduplication codes attenuation of the semantic property of the modifier: *bàbba* ‘big,’ *bàbba-bàbba* ‘biggish’; *ɗanyē* ‘raw,’ *ɗanye-ɗanye* ‘rawish’ (Jaggar 2001).

Some languages show evidence of reduplication in nominal stems, though the source may not be evident: Masa *kēbēkēbe* ‘base of the wing,’ *kijérjēr* ‘morsel’; Giziga *bòbògòf* ‘lungs,’ *gàgàmáy* ‘cotton,’ *kisikid* ‘calabash.’

Reduplication is widely attested in ideophones, a common category across Chadic languages and in many other languages of Africa (see Section 32.4.2). While some ideophones are clearly onomatopoeic, others have no plausible onomatopoeic source, for example, Lele *pílápílá* ‘at the same time,’ *gàgàgàgà* ‘for a long time’ (Frajzyngier 2001: 165); East Dangla *t’árt’ár* ‘quickly,’ *táktákták* ‘completely’ (Shay, field notes).

In Mokilko (East) some ideophones appear to be derived from verbs via reduplication. The reduplicated ideophone is used in a construction with the verb from which it is derived (Jungraithmayr 1990: 105): *gózò-gózòk*, used with *gòzzigè* ‘shake, throw’; *mákà-mákàl*, used with *mákílsà* ‘boast’. The language also has ideophones derived from sources other than the accompanying verb, for example *kórè-kórè-kórèc*, used with the verb *hàrké* ‘frolic.’

In East Dangla, an ideophone derived through reduplication can function as a verb or an adjectival predicate without additional marking: *ɲù sánám sánám* ‘they ate’ (*ɲù* 3PL); *ɲù kirín-kirín* ‘they are fine’ (Shay 1999).

Numerals in some Chadic languages show evidence of reduplication. In many languages, patterns of reduplication suggest an earlier quaternary numeral system (see Frajzyngier and Shay 2012). The evidence is that the numeral “eight” very often represents reduplication of the numeral “four.” For example, the numeral *bídibidów* ‘eight’ in Pero (West) is a reduplicated form of *béedōw* ‘four,’ which itself may result from gemination within the form *bélōw* ‘two’ (Frajzyngier 1989). The corresponding forms in Miya (West) are *ǰādā* ‘four’ and *ǰáǰādā* ‘eight.’ Phonologically similar forms are found in Central Chadic, for example Gidar *pódō* ‘four,’ *dòdòpórò* ‘eight,’ and in East Chadic, e.g. East Dangla *poodi* ‘four,’ *pódópód* ‘eight’; Bidiya *paadan* ‘four,’ *porpod* ‘eight.’

32.8 SUMMARY

Languages of all families of the Afroasiatic phylum show evidence of derivation of nouns from verbs and from other nouns. Languages from all families also exhibit nouns formed via the prefix *m(V)-*, which has been claimed as a characteristic of proto-Afroasiatic and which appears to be productive throughout the phylum, with the exception of the East Chadic and Omotic languages. Derivational prefixation is widespread in Semitic, Egyptian, and Berber, while suffixation is much more common than prefixation in Cushitic, Omotic, and Chadic. In languages where adjectives and adverbs can be identified as distinct lexical categories, these categories may be derived from nouns and verbs via reduplication or suffixation. Compounding is attested in all families but is not common. Languages of all families have means of deriving diminutives from other nouns, sometimes through affixation and sometimes through changes in gender.

Chadic languages exhibit all of the derivational means found in other Afroasiatic languages, namely prefixation; suffixation, often accompanied by vowel and tone changes; reduplication; and compounding. As is the case in other languages of the phylum, derivation from verb to noun is more common than derivation from noun to verb. As pointed out in Frajzyngier (2012), all of the derivational means and paths described for Afroasiatic are also attested in non-Afroasiatic language families of Africa, where the derivational means may carry the same functions that they carry in Afroasiatic. It is therefore difficult to claim that any given derivational marker, with the possible exception of the prefix *m(V)-* (Sections 32.3.1 and 32.7.1), is a characteristic of the phylum, as opposed to an areal feature or a language universal.

ACKNOWLEDGMENTS

I am very grateful to Zygmunt Frajzyngier for his critical reading of this chapter and his very helpful comments. I am also grateful to all the authors who contributed to Frajzyngier and Shay 2012, as their work was vital to this chapter. Any mistakes of fact or interpretation are my own.

CHAPTER 33

NILO-SAHARAN

GERRIT J. DIMMENDAAL

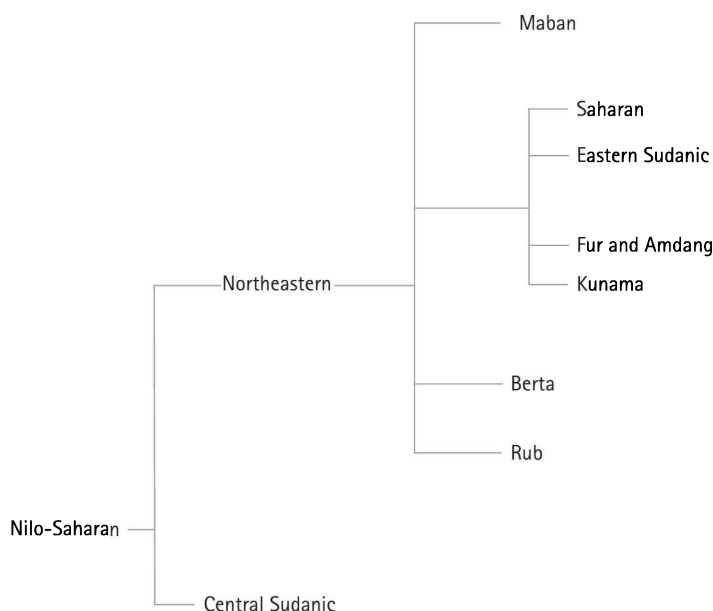
33.1 INTRODUCTION

THE core of the language phylum referred to as Nilo-Saharan today was established more than fifty years ago by the linguist and anthropologist Joseph H. Greenberg. In his earliest genetic classification of African languages, Greenberg (1955) proposed a Macro-Sudanic family (renamed Chari-Nile after a suggestion by the linguist William E. Welmers), consisting of a Central Sudanic and an Eastern Sudanic branch plus two isolated members, Berta and Kunama. This family formed the core of the Nilo-Saharan phylum as hypothesized by Greenberg (1963c). Based on a judicious evaluation in particular of morphological evidence, Greenberg added a number of groups which had been treated as isolates in his 1955 classification, namely Songhay, Saharan, Maban and Mimi, Nyangian, Temainian, Koman, and Gumuz. The present author, however, prefers to treat the Songhay cluster (spoken mainly along the Niger River in West Africa) or Gumuz and the Koman languages (in the border area between Ethiopia and Sudan) as independent families until more extensive grammatical evidence becomes available (Dimmendaal 2011). This also applies to the Kadu languages spoken along the southern range of the Nuba Mountains, Sudan, which had been classified as Kordofanian, that is, as members of the Niger-Congo phylum by Greenberg (1963c), but which have been argued by Bender (1996) to constitute a branch of the Nilo-Saharan phylum too; Ehret (2001: 68) assumes that Kadu as well as the Shabo language in southwestern Ethiopia (usually treated as a linguistic isolate). “may possibly be related at some deeper remove to Nilo-Saharan as a whole.” Scholars thus disagree on the subgrouping or the inclusion or exclusion of specific language families. For further details, the interested reader is referred to the original sources.

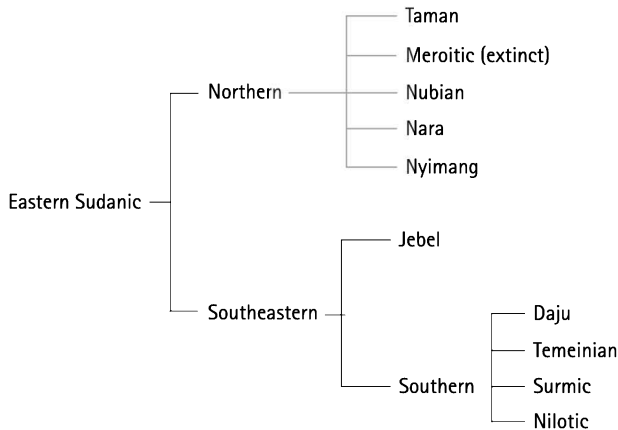
The approximately 120 languages treated as members of the Nilo-Saharan phylum in the present contribution cover an area stretching from Eritrea and Ethiopia in the

East across Northeastern Africa towards Niger with a southward extension into Eastern and Central Africa. Major Nilo-Saharan languages in terms of number of speakers are Dinka (a language in South Sudan which belongs to the Nilotic branch within Eastern Sudanic), Kanuri (a Saharan language spoken in Nigeria and Niger), and Luo (a Nilotic language spoken in Kenya, Tanzania, and Uganda), all of which have millions of speakers.

Dimmendaal (2010) divides the Nilo-Saharan phylum into a Central Sudanic branch and Northeastern Nilo-Saharan, the latter consisting of several well-defined sub-branches:



This primary division within Nilo-Saharan also corresponds to a typological split both syntactically and morphologically. For example, in terms of constituent order Central Sudanic languages have either SVO or SAUXOV constituent order. Within Northeastern Nilo-Saharan, a verb-final syntax is common, with the exception of a group of languages belonging to the Eastern Sudanic subgroup and Berta, which is spoken in the same area. Whereas Central Sudanic languages are predominantly prefixing in their derivational morphology, Northeastern Nilo-Saharan languages tend to use suffixation as a strategy. The largest sub-group, both in terms of number of languages and geographical spreading, is formed by the Eastern Sudanic branch. Within the latter, there is an interesting syntactic and morphological dichotomy between the Northern group and the Southeastern group.



Whereas Eastern Sudanic languages belonging to the Northern branch are strongly dependent-marking at the clausal level (with extensive case marking), Southern members are strongly head-marking (with either a reduced case-marking system or no case marking at all). This typological divergence has had interesting consequences for the lexical, and thereby for the derivational, structure of these languages, as argued in Section 33.5.

33.2 FORMAL STRATEGIES

One of the striking properties of the Nilo-Saharan phylum as a genetic grouping is its typological diversity, a feature related presumably to its geographical spreading and the presence of typologically highly varying neighboring language families such as Cushitic and Chadic, or Ubangian as well as the time depth involved for this genetic unit. With respect to lexical strategies, one may find derivation being expressed by way of tonal changes on words consisting of one consonant or a cluster of consonants only, as in the Central Sudanic language Lendu (spoken in eastern Congo; data with the author):

- (1) b̀r ‘jump’ b̄r ‘jump continuously (pluractional form)’

On the other hand, one finds languages with opulent morphologies, for example within the Nilotic branch of Eastern Sudanic, bordering on Bantu (Niger-Congo) languages, which are also characterized by rich morphologies. In the Nilotic language Kipsikiis (in Kenya), for example, a verb root may be followed by a range of derivational suffixes, themselves followed by inflectional suffixes and clitics. The verb root *-twaal* ‘jump’ may be turned into a causative form by way of the (widespread

Nilo-Saharan) causative prefix *i-*, rendering *-i-twaal*. This latter stem, causing fusion with the infinitive prefix *kee-* (*kee-i-* > *kii-*) if the latter occurs, may be followed by combinations of valency-changing suffixes (example adapted from Toweett 1979: 138):

- (2) kiltwààl-si:-chin-èè-chí lààkwéét
 INF.CAUS.jump-DETR-DAT-MED-DAT child:ABS
 ‘to cause the child to jump for him/her on behalf of somebody else’

Derivation as a lexical strategy is also less prominent in Central Sudanic than in Northeastern Nilo-Saharan; instead, compounding is more common in this primary branch of Nilo-Saharan. Verbal derivation in Central Sudanic languages involves prefixation (next to internal morphology), whereas in languages belonging to the Northeastern branch suffixation constitutes the dominant pattern, as pointed out in the previous section. There is one noticeable exception to this pattern: the prefixation of the above-mentioned causative marker consisting of a high front vowel. This archaic marker is reconstructable for the earliest stages of Nilo-Saharan (Dimmendaal 1982). In Central Sudanic languages, this prefix is part of a set of valency-changing prefixes, as first described in Tucker (1967: 204–19). Blackings and Fabb (2003: 76) give the following examples for Madi (spoken in South Sudan and Northern Uganda):

- (3) tū ‘to climb up’ ī-tú ‘to make climb up; promote’

Prototypically, the prefixation of the causative marker results in the addition of an Agent role to an intransitive verb; alternatively, a noun or adjective may form the lexical basis for a derived verb. The causative prefix has been retained as such in different sub-groups of the other primary branch of Nilo-Saharan, Northeastern Nilo-Saharan, for example in Nilotic and Surmic (which belong to the Eastern Sudanic branch), as shown by the following example from the Nilotic language Kipsikiis (Toweett 1979: 117):

- (4) -nér ‘be fat’ -i-néêr ‘fatten’

Unseth (1998: 116) presents a survey of this archaic feature in Surmic languages; an example from Majang:

- (5) -paak- ‘be hot’ -i-paak ‘heat (transitive)’

Central Sudanic languages have a set of verbal prefixes often affecting the valency of verbs, although in several of them this phenomenon is strongly lexicalized. Again in Madi, one finds a ventive (directional) marker *e-*, or an iterative prefix *u-* or *a-*. In Central Sudanic languages such as Ngiti, such derivational prefixes express direction, causation, but also single action and diminutive (i.e. an action which is carried out for a brief period of time, “a little bit” or is just beginning; Kutsch Lojenga 1994: 299). As

the following examples show, these processes also tend to be accompanied by tonal modifications:

- (6) u-drùta 'to uproot, pull out (single action)'
o-drùta 'to uproot, pull out (basic form)'
u-drùta 'to uproot, pull out (pluractional)'
i-drùtá 'to uproot, pull out (diminutive)'

In addition to these vocalic prefixes, there are also CV prefixes in Central Sudanic, for example the widespread pluractional marker tV-, as in Sara Mbay (Keegan 1997: 40):

- (7) ndá 'hit' tí-ndá 'hit many times'

Pluractionality marking on the verb as such is widespread in at least three African phyla, Nilo-Saharan, Niger-Congo, and Afroasiatic. Prototypically, it expresses an event performed several times by a singular subject in the case of intransitive predications, whereas with transitive predications the event affects multiple objects or the object multiple times.

One prominent prosodic feature of Nilo-Saharan as well as Niger-Congo plus neighboring Chadic, Omotic, and Cushitic (i.e. Afroasiatic) languages, apart from tone, is vowel harmony based on the manipulation of the tongue root, hence "Advanced Tongue Root" or ATR harmony. In a classical system, as in the Nilotic language Kipsikiis referred to above, this system involves a set of five [-ATR] vowels, i, ε, a, ɔ, u, and five [+ATR] vowels ɪ, e, ə, o, u, co-occurring in one word, although there may be irregularities because of specific affixes or clitical elements that do not participate in the harmony process. [+ATR] morphemes tend to be dominant in most cases, as shown in the following nominalizations from Kipsikiis (Toweett 1979: 422):¹

- (8) -kér 'shut' kèèr-ìn 'expert shutter'
-kéér 'see' kèèr-ìn 'expert seer'

The degree of segmental reduction and internal morphology is relatively high in the Central Sudanic branch, but a certain degree of "vertical morphology" can be observed in Northeastern Nilo-Saharan languages too. In the Nilotic language Pàri, for example, consonant alternation is a prominent exponent of derivational (as well as inflectional) morphology, as shown in (9) by the alternations between verbal roots and corresponding pluractional forms (data derived from Andersen 1988):

¹ See Casali (2008) for a detailed survey of this areal feature.

(9)	jap-	jaamm-	'open'
	lot-	lɔɔnn-	'roast'
	ɲɔɔp-	ɲɔɔpp-	'make beer'
	tiet-	teett-	'carry on head'

In another Nilotic language belonging to the same (Western Nilotic) sub-group, Dinka, internal morphology has been taken to an extreme, affecting consonants, tone, as well as vocalic qualities in stems, as shown by Andersen (1992–4). Derived stems in Dinka are essentially monosyllabic, but still manifest different layers of derivational and inflectional morphology formally expressed through modifications in vowel length, voice quality (creaky versus breathy voice).² Andersen (1992–4: 61) shows that the morphological layers are simultaneous but “vertically” ordered, with the root as the “deepest” layer, optionally followed by a derivational layer, and followed by an inflectional layer. For example, the verbal suffix *-u* (expressing ventive or movement towards the deictic centre) in Nilotic languages like Kipsikiis is expressed as a prosodic feature “breathiness” on the root vowel, changing creaky vowels to corresponding breathy vowels in Dinka:

(10)	à-wèec	'(s)he is kicking it'
	D-kick:3SG	
	à-wèec	'(s)he is kicking it hither'
	D-kick:VEN:3SG	

Similarly, the itive suffix *-ta* (marking movement away from the deictic centre) in Nilotic languages like Kipsikiis is realized by way of tonal changes on the root in Dinka.

(11)	à-wèec	'(s)he is kicking it thither'
	D-kick:IT:3SG	

The common Nilotic Dative marker **-ki* (Dimmendaal 2009) again is realized by way of tonal changes and a change of creaky root vowels into breathy vowels in Dinka verb stems:

(12)	à-wèec	mòc	'(s)he is kicking it for the man'
	D-kick:B:3SG	man	
(13)	à-wèec	mòc	'(s)he is kicking for him'
	D-kick:B:AP:3SG	man	

Internal morphology involving tonal modification occurs, not only in Nilo-Saharan languages with strongly reduced segmental structures, but also in languages with opulent

² The seven creaky vowels of Dinka probably go back to five [–ATR] vowels, whereas the corresponding breathy vowels go back to [+ATR] vowels (Andersen 1990).

segmental morphology, as in the Saharan language Kanuri, where abstract nouns may be derived from basic nouns by way of a fixed tone pattern low...high (Hutchison 1981: 64–5).

(14)	fèrò	‘girl’	fèró	‘girlish’
	kámú	‘woman’	kàmú	‘typical of women’
	kànùrí	‘Kanuri people’	kànùrí	‘the Kanuri language’
	fàrànsà	‘French people’	fàrànsá	‘French’

Partial or complete reduplication in Nilo-Saharan expresses inflectional properties like plurality (with nouns and adjectives), but also derivational relations as with adjectives or adverbs derived from nouns. This strategy is particularly common in Central Sudanic, as in Avokaya (data with the author):

(15)	àcí	‘fire’	àcíàcí	‘hot’

Alternatively, partial or complete reduplication may express intensity or repetition, as in the Saharan language Kanuri (data from Hutchison 1981):

(16)	kám-	‘cut’	kàkám-	‘keep on cutting, cutting many’

Such widespread patterns of reduplication are not necessarily indicative of a common origin, but instead appear to have emerged independently again and again in individual languages due to the high degree of diagrammatic iconicity involved.

Metrical structures probably constitute a somewhat neglected formal property of African tone languages. Their nature and impact on morphology also needs much further research. Mora counting plays a role, for example, in the Teso-Turkana cluster within Nilotic both with derivational and inflectional processes. The kind of derivational suffix a derived abstract noun takes is determined by the canonical shape of the adjectival or stative verb root (Dimmendaal 1983: 50–1, 223–58, 270–4). Partially reduplicated roots take a nominal suffix *-(u)*, whereas non-reduplicated *-CVCVC* roots copy the last vowel and consonant from the root; *a-* is a gender marker used for feminine nouns in the singular.³

(17)	-jók	‘good’	á-jók-ìs(ì)	‘kindness’
	-rén	‘red’	a-rén-ìs(ì)	‘redness’
	-nónók	‘soft’	á-nónók-(ù)	‘flexibility, softness’
	-lílím	‘cold’	á-lílím-(ù)	‘coldness’
	-pálór	‘bald’	á-pálór-ór	‘baldness’
	-cókól	‘deep’	á-cókól-ól	‘depth’

³ The parentheses indicate that these vowels in Turkana are de-voiced before a pause.

Similar rules involving mora counting apply to inflectional processes, for example to number marking with nouns in Turkana. Monomoraic nominal roots (i.e. roots containing one vowel), almost always take the number suffix *-m/-in* (the alternation being determined by vowel harmony rules in the language); bimoraic roots on the other hand tend to take the number suffix *-ɪ/-i* (depending on vowel harmony rules), whereas with bimoraic nominal roots ending in a consonant, the number-marking suffix tends to be *-a/-ɔ/-o* (whereby the alternation itself again is determined by vowel harmony rules in the language).

(18)

Singular	Plural	
é-cóm	ɲí-cóm-ín	‘baboon’
é-pém	ɲí-pém-ín	‘bed’
á-pòó	ɲá-póó-í	‘hare’
á-wúnó	ɲá-wúnó-í	‘rope’
á-kínán	ɲá-kìnàn-á	‘crocodile’
á-ɲátùɲ	ɲí-ɲátùɲ-ó	‘lion’
á-tápén	ɲá-tàpèn-ó	‘guineafowl’

The formal realization of pluractionality marking in Nilotic languages like Anywa is conditioned by the presence of a long vowel (or diphthong) as against a short vowel in the root preceding the frequentative marker. The segmental reduction process in Pre-Anywa and the emerging syllable structure was bounded and restricted by higher prosodic structures in the language, the metrical foot (see Dimmendaal 2012 for further details).

Verb root		Pluractional marker		Resulting verb stem
C ₁ VC ₂ -	+	*-CVN	→	C ₁ VVN ₂ N ₂ -
C ₁ VVC ₂ -	+	*-CVN	→	C ₁ VVC ₂ C ₂

Examples showing this synchronic alternation with short root vowels in Anywa:

(19)

root	pluractional stem	
cac-	caɲɲ-	‘look for’
cak-	caɲɲ-	‘name’

With roots consisting of a long vowel or diphthong on the other hand, the resulting pluractional stem consists of a long vowel plus geminate obstruent, rather than a geminate nasal.

(20)

tiet-	teett-	‘speak’
waac-	waacc-	‘speak’

33.3 DERIVATION WITHOUT CATEGORY SHIFT

Derivation without category shift is common across the entire phylum mainly with respect to verbs. One other type not involving category shift, noun-to-noun derivation, is essentially restricted to the Saharan branch, and probably relates to the lexical structure of languages belonging to this Nilo-Saharan branch, as argued in Section 33.5. Other lexically complex forms not resulting from category shift tend to involve compounding in Nilo-Saharan. This latter process is common with all major categories such as nouns, adjectives, or verbs.

Apart from the causative, pluractionality is commonly expressed on the verb in Nilo-Saharan languages. Although the formal strategies for pluractionality marking do not necessarily involve cognate morphemes across this phylum, the actual lexical strategy itself is rather stable. In Central Sudanic languages, it is common to find a prefixal strategy, as in Balese (data translated from Vorbichler 1965: 102):

- (21) à-há ‘cut (one)’ ò-há ‘cut (many)’

There is a common tendency towards internal morphology for pluractional marking as a derivational strategy, but also for the development of suppletive basic vs. pluractional verb stems, as in the Central Sudanic language Ngiti (Kutsch Lojenga 1994: 283).⁴

(22)	Singular	Pluractional	
	ubita	ubita	‘to trap (animals)’
	aràta	owuta	‘go’

In languages belonging to the Maban group, we find pluractional as well as singulative marking on the verb. Moreover, this derivational system interacts with a complex and rich system of number marking on nouns (involving singulatives, plurals, replacement, distributives, and transnumeral forms) functioning as subjects or objects of clauses; see Weiss (2009) for further details.⁵

- (23) à-wá:n-á à-wà:k-ír-i
 ‘I poured a bit’ ‘I poured a lot’

⁴ Nilo-Saharan languages of course are not unique in this respect. Aikhenvald (2012: 154–5) describes similar phenomena in Amazonian languages.

⁵ Replacement refers to an inflectional pattern whereby both singular and plural are marked for number, whereas transnumeral forms are neutral with respect to number, i.e. a transnumeral noun can refer to a singular noun or to a set of more than one entity, i.e. a plural noun.

Apart from causative and pluractional marking, directionality marking on the verb, in particular movement towards the deictic centre (the speaker), is very common across the phylum. This marker is usually referred to as the ventive in the study of Nilo-Saharan languages, where it sometimes is in paradigmatic contrast with an itive marker, expressing movement away from the deictic centre. Directionality marking again is expressed by way of prefixes in Central Sudanic and by way of suffixes in Northeastern Nilo-Saharan languages. Blackings and Fabb (2003: 73–82) describe a ventive prefix ε -/e-, for Mādi:

(24)	̄-jā	kî	lîjá	rá
	3-eat	pl	food	AFF
	‘they did eat food’			
(25)	ē-já	kî	lîjá	rá
	(3)-VEN-eat	PL	food	AFF
	‘they did eat food (away from here)’			

As pointed out by Blackings and Fabb (2003: 73), the inferred meaning or conversational implicature in example (25) as against example (24) is that the persons referred to are present with the speaker at the moment of speaking, but the eating took place elsewhere. The ventive consequently expresses the fact that the speaker observed the event with his own eyes. The same verbal strategy is found not only in various other Nilo-Saharan languages, but also in neighboring Niger-Congo and Afroasiatic languages. In Dimmendaal (2014), it is argued that ventive marking is part of evidentiality marking strategies in languages in the area.

Parallel to pluractional marking, directionality may also be expressed by way of suppletive lexical roots, as in Fur (Waag 2010: 196–210):

(26)	jāŋa	‘go to (place of speaker)’	wǎŋa	‘come back to (place of speaker)’
	ʔéla	‘come to (place of speaker)’	píní	‘go back to (place of speaker)’

As shown by the following example from the Nilotic language Maasai (based on Tucker and Mpaayei 1955), two distinct events (“seeing” and “coming”) involving two distinct actors (a third-person singular moving and an unidentified actor observing the movement) may be expressed in one and the same verb, a phenomenon called “alloying” in Alamin et al. (2012). This “lexical fusion” strategy, which is widespread in languages with ventive marking in the area, is presumably an instantiation of economy principles in languages (expressing a macro event involving two separate events in one word).

- (27) ε -dol-un-i
 3-see-VEN-PASS
 ‘(s)he is seen coming this way’

The example from Maasai shows a further common strategy associated with the passive formation in Nilo-Saharan. Passives in this phylum tend to be expressed as impersonal active constructions, i.e. as constructions without a syntactic subject (agent) and without (syntactic) “object-to-subject” raising, but with a formal marker on the verb.

Apart from causatives or passives as valency-changing derivational operations, middle voice marking is usually marked on the verb in Nilo-Saharan. The modification in valency again tends to be expressed by way of suffixes in Northeastern Nilo-Saharan and prefixes in Central Sudanic. An example from Lese (Tucker and Bryan 1966: 35):

- (28) legbwi ‘to change’ lagbwi ‘to be changed’
 change change:MI

Example (29) is from the Northeastern Nilo-Saharan language Fur (Waag 2010: 187), showing the suffixing strategy of this primary branch:

- (29) biris lup-ɔ ‘(s) laid out the mat’
 mat lay-COMPL
 biris lúp-úŋ-ɔ ‘the mat is laid out (by itself or by someone)’
 mat lay-MI-COMPL

33.4 DERIVATION WITH CATEGORY SHIFT

The categorical distinction between nouns and verbs as major categories in Nilo-Saharan languages is fairly clear from a descriptive point of view, and usually does not pose an analytical problem. Other major categories, such as adjectives or adverbs, may or may not exist as basic or derived categories; it is not always easy to distinguish between stative verbs and adjectives. Derivational processes involving shifts between these various categories are particularly common in Northeastern Nilo-Saharan. Reh (1996: 144–59) describes a proto-typical system of nominalizations for the Nilotic language Anywa. These involve semantic roles such as agentives, instruments, action nouns, abstract nouns, and other types. From a formal point of view, suffixation and internal change are involved, sometimes combined with prefixation:

(30)	Detransitivized verb	Noun	
	kúró ‘watch over’	kùrì (sg) / kùyyè (pl)	‘field guard’
	ràŋò ‘look’	ràŋí (sg) / ràŋyé (pl)	‘mirror, spectacles’
(31)	càmò ‘eat’	dì-càmì (sg) / dì-càmmè (pl)	‘glutton’
	ràà ‘stick into’	dì-rāG (sg) / dì-rāí (pl)	‘twirling stick’

- Derivational suffixes expressing nominalization appear to be absent from most Central Sudanic languages, where compounding tends to be used as a lexical strategy instead, as in the following Avokaya example:

- Our current state of knowledge does not allow us to determine whether suffixation was lost in Central Sudanic, or whether the suffixes in Northeastern Nilo-Saharan are due to reinterpretations of former lexical roots in a compound as affixes. Such reinterpretations of course are common cross-linguistically. In fact, they are attested synchronically in at least one language belonging to the Nilotic family within Nilo-Saharan, Luo, which is in the process of developing noun classes parallel to neighboring Bantu (Niger-Congo) languages. The Luo community has absorbed many speakers from these neighboring speech communities over the past century or more. Apart from heavy (unadapted) lexical borrowing, shift-induced interference from Bantu accompanying language shift probably resulted in the emergence of derivational prefixes (Dimmendaal 2011: 193–6).

- (34) dhó-lúô 'the Luo language' (< * dhok 'mouth')
 já-lúô 'Luo person' (< *jal 'guest')
 jò-lúô 'Luo people' (< *jol 'guests')

Less well known in the literature, but probably quite common in Nilo-Saharan as well as other language families, is the derivation of verbs from ideophonic adverbs. Many African languages, including many Nilo-Saharan languages, are rich in ideophonic adverbs (or adjectives).⁶ One common strategy in Kanuri (and other Saharan languages) is to combine these with the light verb "say" (-n) to form verbal predications. Example (35) is from Hutchison (1981: 107):

- (35) kól 'sound of beating with stick, pounding with hammer; describes small head'
 kólkól- n-'tap sharply on a hard surface: prod, provoke, urge to irrational act'

This process may be widespread across the family, but only a few case studies are available so far.

It is also common across the family to derive adpositions from nouns, in particular from terms referring to body parts. Such derived adpositions (or adpositional nouns) may be realized as prepositions or as postpositions, depending on the language. Non-derived adpositions also play a role in derivational processes in Nilo-Saharan. There is a widespread tendency to use attributive phrases introduced by a preposition, in particular the widespread and archaic preposition *ki* (plus variants) to derive nouns; this preposition introduces a range of semantic roles, in particular instruments. Creider and Creider (2001: 128–39) give numerous examples for the Nilotic language Nandi, and also show that such phrases frequently lexicalize due to the loss of the corresponding basic form; the prefix -*p*- in the examples below is an archaic genitive marker expressing inalienable possession in Nilotic.

(36)	root		derived form	
	keny	'long ago'	ki-p-keny	'old-timer'
	kaliang	'flies'	ki-p-kaliang'it	'fly whisk'
	(absent)		ki-p-chuseit	'puff-adder'

This tendency to lexicalize headless attributive phrases introduced by prepositions (or relative clause markers) as nouns appears to be a permanent trait of Nilo-Saharan languages. One consequence of this "univerbation" process is the frequent occurrence of "empty morphs" in nouns, reflecting former affixes or clitics which have become incorporated into the noun. Dimmendaal (1983: 252–3) lists five such elements for the Nilotic language Turkana.

⁶ Storch and Vossen (2007) describe such rich inventories related to, for example, odors and colors in Nilotic languages.

A corresponding phrasal strategy deriving verbal predicates consisting of nouns, adjectives or adverbs plus the light verb “say” (or sometimes “do”), frequently involving a cognate verb *n(i)*, can be observed in a range of Northeastern Nilo-Saharan branches, for example Saharan, Northern Eastern Sudanic languages as well as Kunama and Berta. Tucker and Bryan (1966: 330) give an example of such a “coverb” plus “light verb” construction in Berta:

- (37) kere ‘light (noun)’ kere-n ‘dawn (verb)’

Cognate forms for the marker *-n* in Berta can be found across the phylum, for example in Nilotic languages like Acholi, where the marker for “say,” *ni* or *li* (depending on the speaker), is used to introduce secondary predicates or adverbial phrases with ideophonic complements (example from Crazzolara 1938: 141):

- (38) wàŋɛ wóŋó lí riik-riik
 eyes.3SG.POSS move say IDEO
 ‘his eyes move restlessly’

In Northern Eastern Sudanic languages like Tama, coverb plus light verb allocations are extremely common, and may constitute around 50% of the predications expressing events. Hutchison (1981: 104) points out for the Saharan language Kanuri that only 5% of the verbal lexicon involves basic verbs, the remaining 95% of the predications being expressed by way of this alternative strategy.⁷ Kanuri has a few hundred ideophonic adverbs, several of which can also be used as verbal predications when combined with the light verb “say, think,” as with the ideophonic adverb *bàdák* describing a loud heartbeat:

- (39) kàrəgənzə (bàdák) bádák-cin
 heart IDEO IDEO-3SG:say
 ‘his heart is beating (distinctly)’

Hutchison (1981: 102) also points out that such predications “are created every day and their sources are easily identifiable. For other . . . roots however, it is not as easy to discover the original source.” In other words, for several coverb plus light verb constructions in Kanuri the corresponding ideophonic adverb (or adjective) no longer exists synchronically in the language.

Interestingly, the coverb plus light verb frame is also used to incorporate verbs borrowed from Arabic into Nilo-Saharan languages with such constructions in languages with such constructional frames, for example in Fur (Waag 2010: 88), where the (Chadic) Arabic stem *yjarrib* ‘try’ is incorporated this way:

⁷ Dixon (2002) shows on the basis of data from Australian languages that languages may go through cycles, moving between coverb plus light verb constructions as dominant patterns or basic forms instead.

- (40) járriβen piá
try do
‘try’

One important reason for this presumably is the fact that verbal morphology in Fur and other Nilo-Saharan languages in the area is characterized by complex morphophonological alternations many of which probably need to be memorized by speakers, because of their irregular nature.

If Gumuz and Koman can indeed be shown to be members of the Nilo-Saharan phylum on the basis of more extensive grammatical evidence (pending more detailed analyses of these languages), one would have a group of languages with a probably unique derivational system within this phylum: nouns (referring to body parts) which can be incorporated into verbs, as shown by Ahland (2010).

- (41) jó-kʷ
die-head
‘kill!’
fún-ic buna
smell-eye coffee
‘smell the (liquid) coffee’

33.5 MORPHOSYNTACTIC TYPOLOGY AND LEXICAL STRATEGIES

Whereas, in general, derivational systems do not seem to bear a direct relation to the morphosyntactic typology of individual Nilo-Saharan languages, there are nevertheless phenomena in two genetic sub-groups showing such a (historical) correlation. This is illustrated in this section, first, for Saharan and, next, for Eastern Sudanic.

An initial example of a correlation between syntactic and lexical strategies comes from the Saharan branch within Nilo-Saharan. Languages belonging to this branch often have relatively few basic verbs, and instead derive lexemes from a combination of the light verb “say” plus coverb, as pointed out for Kanuri in the previous section. Corresponding verb to noun or light verb plus coverb nominalizations are extremely rare in this language, although a few exist:

- (42) dé- ‘cook, boil’ dè-nó ‘cloth boiled after dying’
kúr- ‘draw, mark’ kùr-nó ‘design, drawing’

Noun to noun derivations on the other hand are common in Kanuri and other Saharan languages. Prefixation of *n-*, *nəm-* occurs in order to derive abstract nouns from adjectives or other nouns in Kanuri (Hutchison 1981: 55, 81 ad passim):

(43)	sótó	‘receive a guest’	kù-sótò	‘guest, stranger, foreigner’
	k-âm	‘person’	n-âm	‘being a person’

Suffixation patterns are more common, as is true for other Northeastern Nilo-Saharan languages. This also applies to nouns derived from other nouns in Kanuri:

(44)	ngàbdò	‘garbage’	ngàbdò-lá	‘garbage heap’
	dòndi	‘sick’	dòndi-lá	‘often sick person’
	àlîn	‘indigo’	àlîn-rám	‘dyer’s work site’
	cíntá	‘wiping’	cíntà-rám	‘duster’
	àdîn	‘religion’	àdîn-má	‘devout person’
	cî	‘mouth’	cì-má	‘ambassador’
	jìrè	‘truth’	jìrè-má	‘honest person’
	méli	‘dung’	méli-mà	‘stableboy’

Jakobi and Crass (2004: 112–20) show that similar derivational strategies involving noun-noun relations are attested in the Saharan language Beria:

(45)	kòū	‘song’	kòū-ṛá	‘singer’
	béri	‘Beri person’	bír-dā	‘Beri woman’
	síndó	‘date’	síndó-rā	‘date palm’
	kírá	‘mother’	kírá-ní	‘parents on mother’s side’
	tàmi	‘lamb’	tàmi-nī	‘young lamb’
	kàri	‘Maba people’	kàri-kí	‘the Maba language’

Verbal derivation in Nilo-Saharan languages commonly involves valency-changing operations such as causative, middle voice, antipassive, or pluractional and ventive marking (as illustrated in Section 33.3), and less commonly the thematic incorporation of peripheral semantic roles like instrument or accompaniment. Datives usually occupy an “in between” position in this semantic hierarchy. In other words, if semantic roles like instrument, accompaniment, or manner are expressed on the verb, the dative (covering semantic roles like benefactive, malefactive, or experiencer) is marked on the verb too. This also applies to the Saharan languages discussed earlier in this chapter. With respect to the Saharan language Kanuri, for example, Hutchison (1981: 133–55) makes reference to an “applied” *-kə-* affix expressing semantic roles such as recipient, beneficiary, but also direction, localization, and extending into accompaniment, imitation, effort, intensity, and causative action (Lukas 1937a: 80–3, Hutchison 1981: 141). Such semantic networks or mappings for derivational affixes of course are more common, but cannot be discussed here for reasons of space.

Dative marking is also found in Northern members of Eastern Sudanic, such as the Nubian languages. Comfort and Jakobi (2011) show that in Uncunwee, for example, this verbal extension derives from the verb “give.” A much more extensive marking of derivational relations on the verb, involving core as well as peripheral semantic roles, is common in the Southern members of the Eastern Sudanic branch within Nilo-Saharan, in particular in Surmic and Nilotic. And here is a second domain where it can be shown that morphosyntactic typology and lexical-derivational strategies are related, at least historically. In the Northern representatives of Eastern Sudanic (in particular in Taman, Nubian, and Nyimang), a verb-final constituent order, extensive case-marking systems for core and peripheral arguments, the frequent use of coverb plus light verb constructions, and valency-changing operations essentially affecting core functions (subject and object) are common. Southern representatives of Eastern Sudanic, such as Nilotic and Surmic, are verb-initial or verb second, with a restricted case-marking system (or no case marking at all), whereby a range of semantic roles expressed through case on dependent categories (i.e. syntactic arguments) in Northern Eastern Sudanic are expressed on the verb instead (Dimmendaal 2005). The following examples from the verb-initial Nilotic language Maasai, which only distinguishes between Nominative and Absolutive case, illustrates this (data from Tucker and Mpaayei 1955):

(46)	a-dót-ú	nkójít	
	1SG-pull-VEN	grass:ABS	
	‘I pull out grass’		
(47)	á-írrág-áá	Náròk	
	1sg-sleep-IT	Narok:ABS	
	‘I sleep at Narok’		
(48)	á-ból-óki	papá	ólbéné
	1SG-open-DAT	father:ABS	basket:ABS
	‘I open the basket for father’		
(49)	á-dún-íé	enkálém	
	1SG-cut-INST	knife:ABS	
	‘I cut it with a knife’		

The head-marked clause pattern favors verb-initial constituent order, as shown in Nichols (1986). Verb-initial Nilotic languages confirm this tendency. Several of the derivational suffixes probably go back to prepositions which became bound morphemes on the verb, for example the preposition *ki* (plus allomorphs), which is well-attested in Central Sudanic as well as Northeastern Nilo-Saharan; Dimmendaal 2009, 2014) and which shows up as a dative marker in Nilotic. Synchronically, however, this head marking on the verb is to be treated as a lexical phenomenon, rather than as a case of “syntactic incorporation.” First, the question of whether a verb can take a particular verbal suffix marking dative, instrument, etc., is lexically determined. Second, Maasai and other Nilotic or Surmic languages with extensive head marking on the verb usually have

lexicalized forms with a dative or other derivational suffix for which the corresponding non-derived root no longer exists. The thematic and formal incorporation of these former prepositions is, accordingly, to be treated as a historical, rather than a synchronic, process.

Once such a marker occurs on the verb, it starts interacting with the inherent meaning of the verb and may develop additional shades of meaning. For example in Turkana, a language which is closely related to Maasai, the dative covers semantic roles such as benefactive, malefactive, location, but also experiencer or shades of meaning expressed by way of the suffix *-ish* in English, as part of the semantic mapping.

(50) Root		Dative stem	
-inɔk	'light'	-inɔk-akɪn	'light for somebody'
-iboy	'stay'	-iboy-ikin	'stay somewhere'
(root absent)		-iw-akɪn	'put somewhere'
-rɛŋ	'be red'	-rɛŋ-ikɪn	'be reddish'

Within the Eastern Sudanic branch of Nilo-Saharan, we thus observe a dramatic divergence between Northern representatives (such as Nubian or Taman) and Southern representatives (such as Nilotic or Surmic) in constituent order and the way the interface between morphosyntax and semantics is structured, a divergence which also had radical consequences for the lexical-derivational structure of these languages. Northern members probably represent the more archaic stage (Dimmendaal 2005), since typologically similar systems (with less extensive head marking on the verb and extensive case marking instead) are common elsewhere in Northeastern Nilo-Saharan, and since remnants of this latter system (e.g. relics of peripheral case marking) can still be found in Nilotic and Surmic.

ACKNOWLEDGMENTS

The final draft of the present contribution was written during a stay at the Cairns Institute, James Cook University (Cairns, Australia). I would to thank Alexandra Y. Aikhenvald and R. M. W. Dixon for their kind hospitality and for making this visit possible.

CHAPTER 34

SINO-TIBETAN

KAREN STEFFEN CHUNG, NATHAN W. HILL, AND
JACKSON T.-S. SUN

ALTHOUGH there is a fairly wide consensus on the existence of a Sino-Tibetan family of languages, there is considerable disagreement among scholars regarding its exact internal structure and the hierarchical relations between its members (Jacques forthcoming: 1, van Driem 2011). Many languages in this family are furthermore not well documented or studied. This chapter will not attempt to address or take a stand on genetic affiliations, but will instead simply outline the derivational morphology of three representative Sino-Tibetan languages: Standard Chinese, standard written or Classical Tibetan, and Rgyalrong, a minority language spoken in Sichuan province, China. Perhaps the most high-profile Sino-Tibetan language besides Chinese and Tibetan is Burmese, the national language of Myanmar, which is, however, not treated in this volume. For detailed information on the morphology of Burmese, see Okell and Allott (2001). For recent overviews of the Sino-Tibetan family, see Thurgood and LaPolla (2003), van Driem (2003), and Handel (2008).

PART 1 STANDARD CHINESE

KAREN STEFFEN CHUNG

34.1 INTRODUCTION

Standard Chinese, also called Mandarin Chinese, is the national language of both the People's Republic of China and Taiwan, and one of the official languages of Singapore. It belongs to the Sinitic sub-family of Sino-Tibetan, which includes all Han Chinese dialects. Standard Chinese is based on the Northern Han dialect spoken in and around Beijing. Ethnologue cites a 2000 census giving the number of speakers as 840,000,000 and increasing.

Typologically, Chinese is perhaps the most representative and also one of the most extreme examples of an analytic, isolating language. It is often classified as “mono-syllabic,” though some disagree with this description (e.g. DeFrancis 1984: 177–188, Duanmu and Zhang 2010), in that a large proportion of the lexical items in use in modern spoken Chinese tend to be polysyllabic, often disyllabic; for example, 帽子 *màozǐ* ‘hat,’ 電腦 *diànnǎo* ‘computer’ (n.); 檢討 *jiǎntǎo* ‘to examine, to take stock of,’ 申請 *shēnqǐng* ‘to apply for’ (v.); 聰明 *cōngmíng* ‘intelligent,’ 辛苦 *xīnkǔ* ‘arduous, with difficulty’ (adj./stative verb/adv.); 雖然 *suīrán* ‘although,’ 但是 *dànshì* ‘but’ (conj.). Compound nouns are the most likely to have more than two syllables: 洗髮精 *xǐfǎjīng/xǐfǎjīng* [to wash + hair + essence] ‘shampoo,’ 百貨公司 *bǎihuògōngsī* [hundred + goods + company] ‘department store,’ 健康證明書 *jiànkāngzhèngmíngshū* [health + proof + document] ‘health certificate’; also some foreign-influenced verbs: 再投資 *zàitóuzī* [again + to put-in + capital] ‘to reinvest.’¹

Y. R. Chao, however, says that the “‘monosyllabic myth’ is in fact one of the truest myths in Chinese mythology” (1968: 139). This is because in the vast majority of cases, each syllable is a morpheme, very often a free morpheme, or a morpheme that is only semi-bound; fully bound morphemes are not that frequent in Chinese. In a 200-character survey, DeFrancis (1984: 184) found that 44% of the total were free morphemes, of which 7% were literary forms; 45% were semi-bound; and only 11% were completely bound. Examples of the relatively small store of completely bound syllables include the [+human] plural suffix 們 *men*, which must be attached to a pronoun or noun, and the clitic particle 著 *zhe*, a progressive aspect marker, which can only occur after a verb.

The status of each syllable as a morpheme has enormous consequences for the structure of the language. Unlike English, and particularly unlike its heavy layer of Romance, Chinese does not have many reduced forms that are a close structural match for bound English affixes, stems, roots, and combining forms. In addition, Chinese has one of the lowest ratios of foreign loans—one source puts it at about 1.2% (Wiebusch and Tadmore 2009: 582)—preferring by far loan translations to phonetic loans. Moreover, the foreign origins of phonetic loans that manage to survive beyond an initial burst of popularity are not always apparent to all speakers, since the phonetically similar syllables chosen will frequently also express the meaning of the original to some extent; for example, a popular fast-food chain’s Chinese rendering of “drive-thru” is 得來速 *déláisù*, [to get + to come (marks motion toward speaker) + fast], or ‘get it fast’ (Mandarin has no interdental fricatives and no consonant clusters). Some Mandarin speakers were totally surprised when they first saw the English source of the expression—they had no idea that it was in fact a phonetic loan.

This explains two key features of Chinese: (1) a high semantic density per syllable; and (2) high transparency, due to its use of mostly native components—including archaisms

¹ The Hanyu Pinyin Romanization system with citation tones is used in all examples; tone sandhi are not indicated.

from classical Chinese—rather than foreign ones in word formation. One index of this higher per-syllable semantic density is speech rate. Mandarin was found by Pelligrino and Marsico (2011: 545) to have a very high per-syllable information density index,.94, and a correspondingly lower average rate of syllables spoken per second, 5.18, compared to Spanish and Japanese, with information density indices of.63 and.49, and rates of syllables per second at 7.82 and 7.94 respectively.

So if the conventional definitions of an affix are strictly adhered to, then Chinese has very few affixes indeed, and a number of the ones that qualify are in fact foreign loans, usually calques. Packard (2000: 174), for example, lists only seven prefixes, and some of these are sometimes used as free morphemes. The application of Western European language-based definitions to Chinese not only results in a highly lopsided system, but it also means that many of the native patterns that occur in Chinese will be passed over and perhaps not even recognized at all. For this reason, it makes sense to first describe the patterns observed to occur in Chinese, next to compare them with Western European-type structures, and then to determine whether some adjustments in the definitions might not be in order to better represent the rich morphological patterns of Chinese on their own terms.

34.2 MORPHOSYNTACTIC FEATURES

Chinese is an uninflected language; it does not mark for person or tense, and only marks gender and number in some instances lexically; but it does have aspect-marking particles. Adjectives also function as stative verbs, that is, they can be sentence predicates without a “be” verb, for example 好神奇! *hǎo shénqí* [very + amazing] ‘That’s amazing!’ It has some, mostly lexical, remnants of honorific and deferential speech. Northern Mandarin has an inclusive “we,” 咱們 *zámen*.

The most common method of word formation is compounding. Attempting to distinguish derivation from compounding is not at all straightforward, and is very much open to varying interpretations and approaches. In Chinese, the two are in fact part of the same continuum, rather than cleanly separable domains.

Rather than relying on word endings or internal vowel changes, key parameters in Chinese morphology and morphosyntax are: (1) morpheme/word order; (2) inseparability; (3) word-internal grammatical relations; and (4) syllable count. These will be discussed in turn below.

(1) **Morpheme/word order:** Assuming a Bloomfieldian definition of morphological “head” as that which is ultimately modified in a subordinative formation or construction, and that belongs to the same part of speech as the whole (Bloomfield 1984: 195), Chinese is a head-final language—modifiers *always* precede what they modify, both in the syntax and in the morphology. The ordering of modifiers in Chinese follows basically the same rules in the morphology as in the syntax, with elements describing more

inherent qualities coming closer to the head, and more alienable ones progressively distanced, similar to English, for example 一隻大黑狗 *yī zhī dàhēigǒu* [one + one-of-a-pair (classifier) + big + black + dog] ‘a big black dog’ (‘black’ is more inalienable than ‘big’). Chinese is also basically an **SVO language**; it is at the same time a “**topic-comment**” language, in which background information comes first, and new information is utterance-final.

(2) **Inseparability**: Because, aside from a small number of particles, there are so few overt grammatical markings of any kind in Chinese, and no spaces are left between words in the Chinese writing system, some other means of distinguishing “words” and other morphological formations from syntactic constructions is required. The most consistent and reliable feature marking a word, as opposed to a phrase or other syntactic formation, is *inseparability*. This clearly excludes VO phrases such as 吃飯 *chīfàn* [to eat + rice] ‘to eat’ and 念書 *nìànshū* [to read + book] ‘to study.’ While Chinese does very often prefer a VO construction for many notions that in English would be expressed with a bare verb, these are clearly syntactic phrases and not compounds, since they are as separable as *eat (some) rice* and *read (a few) books* are in English, and certainly no one would call these “compounds.”

Two other common types of syntactic phrases as opposed to compounds in Chinese (mentioned here because like VO phrases, they are erroneously called “compounds” in some accounts) are resultative constructions such as 吃光 *chīguāng* [to eat + bright/bare] ‘to eat something completely up,’ and directional constructions like 走上去 *zǒushàngu/qu* / *zǒushàngqù* [to walk + up + to go] ‘to walk up.’ These types of constructions are very similar to English phrasal verbs such as *to give up* and *to run away*. Potentiality in Chinese can be expressed with one variety of these phrases, by inserting the particle 得 *de* ‘to get’ for “ability” or 不 *bu* ‘not’ for “inability”; for example, 走得動 *zǒudedòng* [to walk + to get + to move] means ‘able to walk,’ 走不動 *zǒubudòng* [to walk + not + to move] means ‘unable to walk.’ While these constructions are highly cohesive, they are also easily separable, and they thus belong to the syntax, and/or perhaps to a category midway between syntactic phrases and compounds, which we may call “collocations.” But they are not “words,” any more than the phrase *to hurry up* is a single word in English (Chung 2006: 23–5).

(3) **Word-internal grammatical relations**: Since there is mostly no morphosyllabic reduction in Chinese to form truly bound morphemes like English *pre-*, which could be used as one of the benchmarks of affixation, and since boundness does not seem to form a consistent or meaningful natural category in Chinese, the next logical place to look is word-internal grammar. Doing so will help uncover the systematic, native patterns of word formation in Chinese rather than only those structures that resemble what we expect to find, based on currently established models of morphological derivation.

One feature that yields consistent results is that of subordination, that is, identifying modifier–head relationships, as opposed to coordination, in which two elements have equal status and there is therefore no head. This approach will help unveil a widespread

system of something like prefixation, but with many semantically strong “prefixes,” which Chao calls “versatile first morphemes in compounds” (1968: 211). These are closer in nature to Greek and Latin combining forms in English, such as *homo-* and *iso-*. These “first morphemes” may in the case of verbs, for example, add information regarding intensity, manner, or direction of an action, but they are usually not as semantically specific as forms like *cardio-* and *astro-*. Combining forms are in fact thus far quite ill-defined in morphology in general—as are these “prefixal” forms in Chinese—and may tend to be lexical in nature. Yet in Chinese they also exhibit considerable productivity (see Section 34.4.2 on verb prefixation for examples). Data and findings from Chinese could perhaps introduce new approaches for the morphological analysis of English and other languages.

(4) **Syllable count:** There is no set number of syllables required to form a word in any part of speech in isolation, but there are very marked tendencies regarding syllable count, less so for nouns, but very strongly so for verbs, adjectives, adverbs, pronouns, and conjunctions—all of these tend to be either one or two syllables in length. Most classifiers are monosyllabic; there are a very few disyllabic compound classifiers, like 回合 *huíhé* ‘a round [in a fight]’; these are more common in the Beijing dialect than in general standard Mandarin. The most basic, common verbs tend to be monosyllabic, for example 吃 *chī* ‘to eat,’ 跑 *pǎo* ‘to run.’ True compound verbs—not syntactic collocations—in general tend to be disyllabic. The disyllabic template is a tightly integrated and inseparable unit which plays an indispensable, central role in all of Chinese morphology.

A given syllable is often bound, in that, in order to be used at all, it needs to occur in a disyllabic expression. The other syllable may be a modifier, a head, or a postposition—the important thing is that the disyllabic template be filled out. 桌 *zhuō* ‘table,’ for example, does not usually appear alone, but 桌子 *zhuōzi* [table + (nominal particle)] ‘table,’ 桌上 *zhuōshàng* [table + top (postposition)] ‘on the table,’ and 飯桌 *fànzhuō* [cooked grain/rice + table] ‘dining table’ are all fine. As Duanmu and Zhang (2010) put it: “The apparent ‘boundness’ of many morphemes, therefore, [is] not morphological but phonological, the avoidance of a monosyllabic foot.”

In addition, Chinese has a form of “syllable harmony,” in which it is often preferred to match the number of syllables in, for example, the object according to the number of syllables in the verb, or the number of syllables in the adjective or adverb to that of the noun or verb it modifies. The monosyllabic attributive 互 *hù* ‘mutually, each other’ appears in the compound verb 互打 *hùdǎ* [mutually + to hit] ‘to hit each other.’ When a disyllabic verb like 毆打 *ōudǎ* ‘to beat up’ is used, it will often be matched with the disyllabic form of the adverb, 互相 *hùxiāng*, also ‘mutually, each other’; thus 互相毆打 *hùxiāngōudǎ* ‘to beat each other up’ instead of 互毆打 *hùōudǎ*, which is also possible but less preferred.

With the above as a bare-bones understanding of how words are constructed in Chinese, we will now describe some of the key patterns of morphological derivation in Mandarin.

34.3 REDUPLICATION

Reduplication is used extensively in Chinese, and has several different uses and meanings.

34.3.1 Nouns

- (a) Some nouns can be reduplicated to mean “every”: 人人¹rénrén [person + person] ‘everybody’; 時時²shíshí [time + time] ‘at all times, at any time’; this can also be done with some classifiers, e.g. 條條³(道路) tiáotiáo (dàolù) [strip + strip (+ road)] ‘every road’; 層層⁴céngcéng [layer + layer] ‘[at] every level, every layer.’ Reduplication of pairs of related nouns or classifiers can add emphasis: 分分秒秒⁵fēnfēnmiǎomiǎo [minute + minute + second + second] ‘every minute and every second, at all times.’ This pattern may have the extended meaning of ‘many’ or ‘one thing happening after the other, one at a time’: 層層疊疊⁶céngcéngdiédié [layer + layer + fold + fold] ‘having many layers, layer by layer’; 點點滴滴⁷diǎndiǎndīdī [dot + dot + drop + drop] ‘bit by bit, slowly.’
- (b) Some reduplicated nouns are fixed lexical items rather than on-the-spot output of a morphological pattern or rule. Reduplication is sometimes used, for example, in the names of animals: 猩猩⁸xīngxīng ‘gorilla,’ 狒狒⁹fèifèi ‘baboon,’ 螞蟥兒¹⁰guōguor ‘cricket’ (N. Mandarin; this seems to be onomatopoeic; alliteration and rhyme are also common in disyllabic names of plants and animals.) In some cases, lexical reduplication has the meaning of ‘many,’ as in (a), e.g. 星星¹¹xīngxīng ‘stars.’
- (c) Many words for family relationships are reduplicated, and suggest intimacy: 姐姐¹²jiějie ‘elder sister’; 爸爸¹³bàba ‘father.’
- (d) Nouns are reduplicated and form diminutives in baby talk: 車車¹⁴chēchē [car + car] ‘little car’; 鞋鞋¹⁵xiéxié/xiěxié [shoe + shoe] ‘little shoe’; also playfully, in adult speech: 東東¹⁶dōngdōng [east + east] ‘thing, thingie,’ playful variant of 東西¹⁷dōngxī [east + west] ‘thing, object’ in which the first element is reduplicated. d) Some nouns are reduplicated simply to fill out the disyllabic template; this form may in some cases also add a slightly diminutive or tongue-in-cheek flavor to the expression, as in (d): 包包¹⁸bāobāo [package + package] ‘purse, bag’; 叉叉¹⁹chāchā [cross + cross] ‘an X’; 調調²⁰diàodiào [tune + tune] ‘overall style’; 頭頭²¹tóutou [head + head] ‘leader, chief’; 渣渣²²zhāzhā [dregs + dregs] ‘dregs, debris, trash, junk.’ These may take an 兒²³er suffix in Northern Chinese (see Section 34.5.1).

34.3.2 Verbs

Reduplication of verbs usually means “to try to do something, to do something casually or tentatively, not seriously”: 試試看 *shìshìkàn* [to try + to try + to see] ‘try it and see’; 走走 *zǒuzǒu* [to walk + to walk] (or: 走一走 *zǒuyīzǒu* [to walk + one + to walk] ‘to take a leisurely walk, to walk a bit.’ ‘Thank you’ in Mandarin is a reduplication of the verb “to thank,” 謝謝 *xièxiè*.

Disyllabic compound verbs may follow an ABAB pattern: 慶祝慶祝 *qìngzhùqìngzhù* [to celebrate + to celebrate] ‘to celebrate a bit,’ though AABB is used for copulative compound verbs: 吃吃喝喝 *chīchīhēhē* [to eat + to eat + to drink + to drink] ‘to have a good time eating and drinking’; 進進出出 *jìnjìnchūchū* [to enter + to enter + to exit + to exit] ‘to be constantly coming and going’; these are basically the reduplication of two individual monosyllabic verbs.

34.3.3 Adjectives and Adverbs

Reduplication of adjectives and adverbs usually adds a sense of “quite, rather”: 高高的 *gāogāode* [tall + tall + (particle of manner)] ‘rather tall’; 好好的 *hǎohǎode* [good + good (particle of manner)] ‘just fine, properly.’ Disyllabic adjectives tend to follow an AABB pattern of reduplication: 高高興興 *gāogāoxìngxìng* [high + high + excited + excited] ‘quite happy/happily.’ (高興高興 *gāoxìnggāoxìng* also occurs, but in this case it is a **verb**, meaning ‘to experience feelings of pleasure’). This is an example of an expansion through reduplication of a adjective/adverb often used in the non-reduplicated form 高興 *gāoxìng* ‘happy’ as well. But some reduplicated modifiers may be compounded from expressions that do not exist or are uncommon in the simplex form. In this example the reduplication acts as an intensifier and stresses the disparate nature of a referent: 花花綠綠 *huāhuālǜlǜ* [patterned + patterned + green + green] ‘with gaudy patterns and colors’; the non-reduplicated form *花綠 *huālǜ* is not used. There are also ABB-type adjective and adverb formations, for example 冷冰冰 *lěngbīngbīng* [cold + ice + ice] ‘cold as ice’ and 黑漆漆 *hēiqīqī* [black + lacquer + lacquer] ‘pitch black’; and AAB-type modified nouns, such as 毛毛蟲 *máomáochóng* [furry + furry + insect] ‘caterpillar’ and 甜甜圈 *tiántiánquān* [sweet + sweet + circle] ‘doughnut.’

34.3.4 Onomatopoeia

As mentioned in passing in Section 34.3.1 (b), reduplication is frequent in onomatopoeia, for example 哈哈笑 *hāhāxiào* [*hā* + *hā* + to laugh] ‘to laugh out loud’; 嗶嗶叫 *bībījiào* [*bī* + *bī* + to call out] ‘to beep’; 嘰嘰呱呱 *jījīguāguā* ‘to chatter endlessly’; 咔嚓咔嚓 *kāchākāchā* ‘click, click.’

34.4 AFFIXATION: PREFIXATION

Chinese has formations such as 反恐怖主義者 fǎnkǒngbùzhǔyìzhě [anti- + terror + -ism + -er/-ist] ‘anti-terrorist,’ which seem to display a pattern of affixation tidily similar to English—except that words like these are very modern ones, in which the “affixes,” while all formed from native Chinese elements, are in many cases simply loan translations from English or other Western languages, often with Japanese as a middleman. The danger in trying to impose a pre-established system of affixation on Chinese is that the examples that best fit the requirements looked for will be ones that were borrowed from the language on which the classical prefixation model was based in the first place. So the search becomes a circular one, going from an English system to a parallel borrowed English-based system, while more native Chinese patterns with much deeper roots, such as Chao’s “versatile first morphemes in compounds,” are passed over.

34.4.1 Nouns

There is only a small number of widely-recognized noun prefixes in Chinese following the conventional definitions. Even these often do not strictly meet the usual criteria, in that some, such as 老 lǎo ‘old,’ can be used as free morphemes. Here are some examples:

1. 阿 ā: A fully bound prefix for nicknames and certain family relationships; e.g. 阿姨 āyí [ā (noun prefix) + mother’s sister] ‘Auntie’; 阿強 āqiáng ‘A Qiang,’ nickname for someone with “Qiang” as the third character of his name.
2. 老 lǎo ‘old’: This is a marginal noun prefix with bleached meaning used in just a few established nouns; e.g. 老師 lǎoshī [old + teacher] ‘teacher,’ 老虎 lǎohǔ [old + tiger] ‘tiger,’ 老鼠 lǎoshǔ [old + mouse/rat] ‘mouse, rat.’ It is also used as a term of friendly respect before family names and certain family relationships: 老王 lǎowáng ‘Old [Mr.] Wang,’ 老哥哥 lǎogēgē [old + elder brother (reduplicated)] ‘[dear] big brother.’
3. 第 dì: A bound prefix used to form ordinal numbers, e.g. 第一個 dìyīgè [(ordinal number marker) + one + (classifier)] ‘first.’
4. 前 qián ‘former, ex-’ and
5. 副 fù ‘deputy, vice-’; e.g. 前副總統 qiánfùzǒngtǒng [previous + vice- + president] ‘former vice president’—an example with two prefixes used together.
6. 準 zhǔn ‘future, -to-be, quasi-’: 準媽媽 zhǔnmāmā [future + mother] ‘mother-to-be’; 準媳婦 zhǔnxífù [future + daughter-in-law] ‘daughter-in-law-to-be’; 準貨幣 zhǔnhuòbì [quasi- + currency] ‘quasi-money.’
7. 迷你 mí nǐ ‘mini’: An example of a recent disyllabic prefix that is a phonetic foreign loan; e.g. 迷你裙 mínǐqún mí + nǐ + skirt] ‘miniskirt’; 迷你火鍋 mínǐ

huǒguō *mí* + *nǐ* + fire + pot] ‘mini-hotpot, individual chafing dish stew.’ This prefix is now semantically empty in most formations. The two syllables, however, may have originally been chosen to add semantic content to what is probably the earliest loan using them, 迷你裙 *mínǐqún* ‘miniskirt,’ which could be glossed as [to dazzle + you + skirt] ‘skirt that will dazzle/stun you.’

34.4.2 Verbs

There is an extensive system of Chao’s prefix-like “versatile first morphemes” in Chinese that form compound verbs. Some examples:

1. 預 *yù* ‘pre-, beforehand’: 預熱 *yùrè* [beforehand, pre- + to heat] ‘to preheat’; 預付 *yùfù* [beforehand, pre- + to pay] ‘to pay in advance, to prepay.’
2. 複 or 復 *fù* ‘re-, again’: 複查 *fùchá* [again, re- + to examine] ‘to reexamine’; 復發 *fùfā* [again, re- + to erupt] ‘to have a relapse.’
3. 誤 *wù* ‘mis-, mistaken’: 誤導 *wùdǎo* [mistaken, mis- + to lead] ‘to mislead’; 誤診 *wùzhěn* [mistaken, mis- + to diagnose] ‘to misdiagnose.’
4. 合 *hé* ‘co-, together, jointly’: 合作 *hézuò* [together, co- + to do] ‘to cooperate’; 合辦 *hébàn* [jointly + to hold, to administer] ‘to jointly hold (an event)’; 合唱 *héchàng* [together + to sing] ‘to sing in chorus’ (Chung 2006: 33–73).

34.5 AFFIXATION: SUFFIXATION

34.5.1 Nouns

1. Bound plural marker 們 *-men*: This straightforward example of a suffix is added to the three singular pronouns, 我 *wǒ* ‘I’, 你 *nǐ* ‘you’ (the /-n/ of the formal 您 *nín* originated in the plural 們 *-men* form, and is limited in use), 他 (general) / 她 (marked as feminine, ‘she’) / 它 (neuter, ‘it’) / 牠 (for animals) / 祂 (for deities), all pronounced *tā*, ‘he/she/it’, for the corresponding plural forms: 我們 *wǒmen* ‘we, us’, 你們 *nǐmen* ‘you (pl.)’, 他們 *tāmen* ‘they, them.’ It can also be added to [+human] nouns in some cases, e.g. 孩子們 *háizimen* [child + (noun suffix) + (plural marker)] ‘children’; 同學們 *tóngxuémen* [same + to learn + (plural marker)] ‘fellow students’; this form is often used in appeals or calls to action.
2. 子 *zi*: This nominal suffix (original meaning: ‘child’) is added to many nouns or other formations to make them disyllabic or polysyllabic free morphemes, most of which are lexical; e.g. 袋 *dài* ‘bag, pocket’ is mostly a bound morpheme; with the addition of 子 *zi* it becomes the free morpheme 袋子 *dàizi* ‘bag’; 扇子 *shànzi*

[fan + (nominal particle)] ‘fan’; 包子 *bāozi* [package + (nominal particle)] ‘steamed filled bun.’ It is also used to form some tongue-in-cheek nouns such as 左撇子 *zuǒpiězi* [left + left-falling-brush-stroke + (nominal particle)] ‘leftie, southpaw, left-handed person’ and 嘴皮子 *zuǐpízi* [mouth + skin + (nominal particle)] ‘lips (of a glib speaker)’.

3. 頭 *tou* ‘head’ (cf. English ‘to come to a head’, i.e. this suffix suggests the culmination, concentration or crystallization of the essential elements of a thing or notion): 頭 *tou*, like 子 *zi*, creates free morphemes from certain bound nouns, or expresses a concrete aspect of a monosyllabic noun, e.g. 石頭 *shítou* [stone + *tou*] ‘stone’; 鐘頭 *zhōngtou* [clock + *tou*] ‘hour’; it can also nominalize certain verbs, e.g. 念頭 *niàntou* [thought + *tou*] ‘idea, notion’; 玩頭 *wántou* [to play + *tou*] ‘something fun to do.’ It is usually pronounced in the neutral tone, though it may receive its full second tone value in southern and other varieties of Mandarin.
4. 兒 *er*: In the Mandarin spoken in Beijing and environs, a subsyllabic /-r/ suffix is frequently added to content words, especially nouns, but also verbs, modifiers, and even suffixes, like 頭 *tou*. It often changes the meaning of the non-rhoticized form in a subtle way, e.g. 肝 *gān* ‘liver’ refers to the organ in a living being; 肝兒 *gār* refers to liver as an organ meat for food (cf. Spanish *pez* ‘fish [in the water]’ vs. *pescado* ‘fish [out of the water, as food]’). In some cases it makes a free morpheme of a bound one, e.g. the semi-bound morpheme 味 *wèi* ‘flavor,’ which usually appears in a polysyllabic compound like 味道 *wèidào* [flavor + way] ‘taste,’ is often expressed as the monosyllabic 味兒 *wèr* ‘taste’ in Northern Chinese. 兒 *er* may also mark part of speech, e.g. a noun as opposed to a verb, as in 蓋 *gài* ‘to cover’ (v.) and 蓋兒 *gàr* ‘a cover, a lid’ (n.). And it can function as a diminutive or hypocoristic suffix: 寶貝 *bǎobèi* is ‘treasure,’ 寶貝兒 *bǎobèr* is ‘little darling.’ It has high but not unlimited productivity in Northern Chinese. It never adds an additional syllable when used as a suffix, and it often causes a final nasal to be dropped and a vowel change in addition. Suffixal 兒 *er* is not much used in southern and certain other varieties of Mandarin; when written, it is often “incorrectly” read as a separate syllable in these varieties.

As can be seen from the 孩子們 *háizimen* ‘children’ example above, more than one suffix at the end of a word is possible, though not extremely common. More than two suffixes is much rarer; a historical example with three suffixes is 小丫頭子們 *xiǎoyātouzimen* [little + lass + *tou* + *zi* + *men*] ‘young servant girls.’

34.5.2 Agentives

The agentives 者 *zhě* ‘one who...’ and 家 *jiā* ‘person, family’ are highly productive, but somewhat marked, beyond established compounds like 作者 *zuòzhě* [create/write + -er] ‘author’ and 音樂家 *yīnyuèjiā* [music + person] ‘musician.’

34.5.3 Individuation

While Chinese does not usually mark plurality, it does have a form that emphasizes a referent as being composed of individual parts, similar to English *hair/hairs*, *fish/fishes*. To specify a certain quantity of a noun, normally a number is followed by a classifier, then the noun, for example 三張紙 sānzhāngzhǐ [three + sheet + paper] ‘three sheets of paper.’ 紙 zhǐ ‘paper’ by itself is a mass noun suggesting no form or number; however, placing the classifier after the noun yields 紙張 zhǐzhāng [paper + sheet] ‘papers,’ stressing its concrete form in sheets, rather than just the notion of ‘paper’; 一隻船 yīzhīchuán ‘a boat’ becomes 船隻 chuánzhī, a collective of individuated ‘boats.’ These forms are lexically determined and the pattern has limited productivity. Chinese also has a number of mostly lexical duals and plurals, such as 雙親 shuāngqīn [pair + parent] ‘[both] parents’; and 對筆 duìbǐ [pair + pen] ‘paired pen set’; and the literary classifier used as polite form of address 諸位 zhūwèi [all + (honorific measure word)] ‘all of you esteemed people.’

34.5.4 Verbs

The most productive verbal suffix is 化 huà. It was first coined by the Japanese in the 19th century as a translation for English *-ize* (*sanitary/sanitize*), *-ify* (*mystery/mystify*), and *-ate* (*automatic/automate*), and was then later borrowed into Chinese for the same purpose. This set of English suffixes is subject to various morphophonemic constraints; in Chinese, however, the 化 huà suffix has virtually unlimited productivity. For example: 全球化 quánqiúhuà [entire + globe + -ize] ‘to globalize/globalization’; 數位化 shùwèihuà [digital + -ize] ‘digitize’; 西化 xīhuà [west + -ize/-ization] ‘to Westernize’/‘Westernization’; 少子化 shǎozǐhuà [few + children + -ize/-ify] ‘tendency for the birth rate to drop.’

34.6 AFFIXATION: CIRCUMFIXES AND INFIXES

Chinese does not seem to have circumfixes per se, but infixation is found as a marginal phenomenon in certain set expressions, mostly onomatopoeic and colorful, humorous ones, the prime example of an infix being 裡 li, also written 哩 or 里, as in 傻裡傻氣 shǎlishǎqì [silly + li (particle) + silly + air] = ‘rather silly’; 小裡小氣 xiǎolixiǎoqì ‘rather stingy’, and the very common 糊裡糊塗 húlihútú ‘confused, in a jumble’, all of which combine infixation with reduplication in an A li A qì or A li AB pattern. Various sets of numbers can be used as infixes, or in other patterns, to express set conventional meanings, for example ‘seven’ used with ‘eight’, or sometimes ‘eight’ alone, suggests ‘chaos, confusion’: 亂七八糟 luànqībāzāo [chaotic + seven + eight + mess] ‘chaotic, in a jumble’; the 裡 li type form can sometimes be combined with numbers to produce expressions such as 囉裡八嗦 luōlibāsuo [(first half of) to nag + li (particle) + eight + (second half of) to nag] ‘to nag, to carry on tediously’, a further intensification of the more common 囉裡囉嗦 luōliluōsuo [(first half of) to nag + li (particle) + (second half of) to nag], which has the same meaning. Infixation is more common in Northern Mandarin.

34.7 CONVERSION

Conversion is common in Chinese. Most content words can easily be used as nouns, and may be preceded by the possessive enclitic particle 的 *de*, for example 漂亮 *piàoliang* is ‘pretty’; 她的漂亮 *tāde piàoliang* ‘her prettiness.’ Nouns are often used as adjectives, sometimes poetically; one commonly-cited example is 今天的天空多麼希臘 *jīntiān de tiānkōng duōmo Xīlǎ* ‘How ‘Greece’ the sky is today,’ suggesting a sky that is a deep Mediterranean blue. Nouns are less often borrowed directly for use as verbs; incorporation is preferred—see examples below. In a few cases, tone distinguishes nouns from verbs, for example the noun “nail” is pronounced in the first tone 釘 *dīng*, as in 釘子 *dīngzi* ‘a nail,’ while the verb “to nail” is pronounced with the fourth tone 釘 *dìng*, thus the cognate VO phrase 釘釘子 *dìng dīngzi*, ‘to pound in a nail’; the noun for “drill” is pronounced in the fourth tone: 鑽 *zuàn*, as in 電鑽 *diànzuàn* ‘electric drill’; while the verb is pronounced in the first tone, 鑽 *zuān*, as in 鑽洞 *zuāndòng* ‘to drill a hole’; and 處 *chù*, in the fourth tone, is ‘place,’ while 處 *chǔ*, in the third tone, is ‘to be located at/in, to manage.’

34.8 INCORPORATION

Verbs with an incorporated noun may be literal, for example 手洗 *shǒuxǐ* [hand + to wash] ‘to hand wash’; 目睹 *mùdǔ* [eye + to see] ‘to witness with one’s own eyes’; 面議 *miànyì* [face + to discuss] ‘to discuss in person, face-to-face’; or metaphorical, like English ‘to pinpoint,’ ‘to cherry-pick’: 冰釋 *bīngshì* [ice + to release] ‘to melt/dissolve away like ice’; 瓜分 *guāfēn* [melon + to divide] ‘to divide up as though cutting a melon, to divvy up, to share the pie’; 蜂擁 *fēngyōng/yǒng* [bee + to crowd together] ‘to swarm’ (Chung 2006: 74–82).

For more comprehensive coverage of derivational morphology in Chinese, the reader is referred to Chao (1968).

PART 2 TIBETAN

NATHAN W. HILL

34.9 BACKGROUND AND INTRODUCTION

In the 7th century the Tibetan empire expanded across the Tibetan plateau and eventually throughout Central Asia (Beckwith 1993). In 650 writing was introduced as

an administrative exigency of running an empire (N. W. Hill 2010c: 110–12); the earliest extant documents date from a century later. Tibetan linguistic history is conventionally divided between Old Tibetan (11th century and earlier) and Classical Tibetan (later texts). Tibetan boasts a vast literature with a wide variety of genres, and the family of Tibetic languages spoken today is comparable in size and diversity to the Romance languages (Tournadre 2008: 282–3). Most researchers see Tibetan as a member of a language family which also includes Burmese and Chinese; this family is known by names including “Tibeto-Burman,” “Sino-Tibetan,” and “Trans-Himalayan,” of which the last is the most neutral and accurate (cf. van Driem 2012).

34.10 PHONOLOGY AND MORPHOSYNTAX

The Tibetan alphabet distinguishes twenty-nine consonants (k, kh, g, ŋ, c, ch, j, ñ, t, th, d, n, p, ph, b, m, ts, tsh, dz, w, ź, z, ɣ, y, r, l, ś, s, h) and five vowels (a, e, i, o, u); the alphabet is a good, but not perfect match to Old Tibetan phonology (cf. Hill 2010c). Syllables are distinguished with explicit punctuation. The limitation of onset clusters to word initial syllables provides a possible definition for a language specific phonemic word; most lexemes so defined would be disyllabic. Old Tibetan did not have tone and the tonal systems of those modern Tibetic languages that possess them derive transparently from segmental phonology. Tibetan has agglutinative morphology and ergative alignment (Tournadre 1996); it exhibits *Gruppenflexion*, with ten morphological cases (cf. Hill 2012). Tibetan lacks any agreement systems, but verbal suffixes indicate switch reference (Andersen 1987, Zadoks 2000, 2002, Haller 2009). Tibetan verbal inflection is complex, with four verb stems showing a variety of ablaut, stem alternation, prefixes and suffixes (e.g. present *hɛdzin*, past *bzuñ*, future *gzuñ*, imperative *zuñs* ‘take’). However, the various stems derive from roots through the application of a small number of affixes with subsequent sound changes obscuring the system (cf. Coblin 1976, N. W. Hill 2010a). One of the few components of verbal inflection yet to receive full clarification is an ablaut from *-a-* to *-o-* in the past stem of a handful of verbs (e.g. *za*, *zos*, *bzah*, *zo*, ‘eat’). Following the technique of Meillet (1925: 25) to locate inherited morphology among irregularities, Jacques (2010b, 2012b) sees this *-o-* in the past as a fragmentary reflex of an erstwhile **u* third-person object agreement suffix also confronted in Rgyalrong and Kiranti; LaPolla (2012) disagrees. Both authors overlook the existence of this pattern in verbs aside from “eat” (e.g. *hɛñ*, *doñ*, —, ‘disappear’, *hchah*, *hchos*, *hchah*, *hcho* ‘chew, gnaw’, *lañ*, *loñs*, *lañ*, — ‘finish’, cf. Hill 2010a: 89, 148, 279).

Very little is securely known about Tibetan derivational morphology and it is difficult to determine when a purported derivational affix was productive. Here it is not feasible to describe the morphology of any particular *état de langue*, I instead survey the findings of available research on ‘Classical Tibetan’, with occasional reference to Lhasa dialect.

34.11 VOICING ALTERNATION

Many authors point to the existence of pairs of verbs, an intransitive with a voiced initial (e.g. *ḥbo*, *bo*, —, — ‘overflow’) and a transitive that exhibits voice alternation among its stems (e.g. *ḥbo*, *phos*, *dbo*, *phos* ‘pour’). There have been attempts to derive the voiced forms from the voiceless (Pulleyblank 1973: 114–16, Sagart 2006, and Jacques 2012d: 215–16) and vice versa (Bielmeier 1988: 19, Beckwith 1999).

Those who derive the voiced forms from the voiceless point to parallels in cognate languages. For example, Sagart compares Tibetan voicing alternation to voicing alternation in Rgyalrong and Chinese. Rgyalrong has verb pairs in which the transitive member is voiceless and the intransitive member is voiced and pre-nasalized, for example *ka-tfop* ‘to set fire to,’ *kə-ndʒop* ‘to catch fire’; *kə-klek* ‘to wipe off,’ *kə-ŋglek* ‘to fall.’ Paralleling the Rgyalrong case and relying on prenasalization in the pronunciation of Chinese loanwords into Hmong-mien, Sagart reconstructs an *N- prefix in Chinese that derives voiced intransitive verbs from voiceless intransitives (2006: 64), e.g. 夾 *C.k^ʕrep ‘press between,’ 狭 *N-k^ʕrep ‘narrow.’

Almost all authors ignore the fact that in Tibetan this alternation is part of a larger pattern involving three stems, a voiced intransitive, a voice alternating transitive, and a voiceless transitive (cf. Uray 1953a: 50–1), e.g. √gab (*gab*, *gab*, *gab*, *gob*) ‘hide intr.’, √gab (*ḥgebs*, *bkab*, *dgab*, *khob*) ‘cover tr.’, √kebs (*khebs*, *khebs*) ‘be covered over.’ The voiceless intransitive appears to derive from the present stem of the voice alternating transitive (cf. Hill 2014). This suggests that the voiced root is primary.

34.12 REDUPLICATION

Reduplication is used in Tibetan for a variety of functions including emphasis (e.g. *sems bloṅ-bloṅ-po* ‘depressed mind,’ cf. *bloṅ-po* ‘dull person’ Uray 1955: 182), onomatopoeia (e.g. *khruṅ-khruṅ* ‘crane, stork’), temporal repetition (e.g. *phyir-phyir* ‘one after another,’ cf. *phyir* ‘back, behind’ Uray 1955: 190), and spatial diversity (e.g. *phyogs-phyogs-su* ‘to different quarters,’ cf. *phyogs* ‘direction’ cf. Uray 1955: 191). Reduplication is more characteristic of the modern languages than of Old or Classical Tibetan. Uray (1955) gives a thorough treatment of phenomena in the published dialectology of his day. Robin and Vittrant (2007) discuss reduplication in Lhasa dialect. Zhang (1985) examines one type of reduplicated word in Lhasa Tibetan in detail.²

² The only common usage in Classical Tibetan is seen not in nominal derivation but in verbal inflection: a reduplicated nominalized verb followed by *-las* indicates an interrupted action, e.g. *de-nas bya-ṇaṅ-pa kun-tu ḥphur-ṇiṅ tshol-tshol-ba-las / skyed-mos-tshal de-na rgyal-buḥi skad grags-nas* ‘While the duck, flying everywhere, was searching [for him], the voice of the prince was heard in the park’ (Uray 1955: 195, cf. N. W. Hill 2010b: 255–7).

34.13 COMPOUNDING

Tibetan exhibits a healthy variety of compounds (Beyer 1992: 104, Hahn 1997, Volkart 2003); examples exist to fit most of the categories known from Sanskrit (Dvandvas: *ra-lug* ‘goats and sheep,’ Karmadhāraya: *dkar-khuñ* ‘window’ lit. ‘white-hole,’ Dvigu: *gsuñs-ḥbum* ‘collected works’ lit. ‘word-100,000,’ Tatpuruṣa: *sbrañ-rtsi* ‘honey’ lit. ‘bee-juice,’ Bahuvrīhi: *dri-bzañ* ‘saffron’ lit. ‘smell-good’). Disyllabic words are often shortened to monosyllables when they enter into compound (e.g. *yi-ge* ‘letter, writing,’ *chen-po* ‘big,’ *lag-pa* ‘hand’ vs. *bod-yig* ‘Tibetan written language,’ *blon-chen* ‘prime minister,’ *lag-rtsal* ‘handicraft’). Some elements are so frequent and of such generalized meaning in compounds as to be identified as suffixes (e.g. *bdag-po* ‘lord,’ *sa-bdag* ‘local spirit,’ lit. ‘earth-lord,’ *sbyin-bdag* ‘patron,’ lit. ‘give-lord’ cf. Beyer 1992: 120). The noun *mkhan-po* ‘abbot, expert,’ seen in *śiñ-mkhan* ‘carpenter,’ and *bzo-mkhan* ‘artisan,’ becomes a productive agentive suffix in Lhasa Tibetan that even operates at the clause level (Denwood 1999: 229).

Honorific body part terminology is used in compounds as modifiers to form honorifics. Thus, *skra* ‘head hair’ in compound with *dbu* ‘head (hon.)’ becomes *dbu-skra* ‘head hair (hon.)’ and *zal* ‘mouth (hon.)’ in compound with *dkar-yol* ‘teacup’ becomes *zal-dkar* ‘teacup (hon.)’ (Beyer 1992: 156–7; Denwood 1999: 95–8). This strategy for forming honorifics is common in the Classical language and remains productive in Lhasa Tibetan. The choice of honorific prefix is highly lexicalized.

34.14 AFFIXATION

Most of Tibetan derivational morphology is affixal; there is a causative denominative prefix *s-* in the verbal system and a variety of suffixes occur in nominal derivation.

34.14.1 Causative and Denominative *s-* Prefix

The most conspicuous derivational affix in the Tibetan verbal system is a causative and denominative *s-* prefix (Conrady 1896). A non-causative verb is paired with a causative verb; the two verbs share a root, but the causative verb is prefixed by *s-* (e.g. *lañ* ‘rise,’ *slañ* ‘raise,’ cf. Chang 1971, Beyer 1992: 116–17, Gyurmé 1992: 258–61). Before voiced laterals *s-* appears orthographically as *z-* (e.g. *lo* ‘report n.,’ *zlo* ‘report v., say,’ cf. Hahn 1999: 124). In some cases *r-* appears in place of *s-* (e.g. *ñil* ‘decay,’ *rñil* ‘destroy’); the reasons for this variation are not understood. There is also a complication in verbs with root initial *b-*; while normally the prefix *s-* does not affect the initial *b-* (e.g. *√byañ* ‘be cleansed,’ *√sbyañ* ‘cleanse’) in rare cases the causative member of the pair has *sp-* rather than *sb-* (e.g. *√bub* ‘be upside down,’ *√spub* ‘invert’). Voicing alternation interacts with

the *s-* prefix to sometimes yield a wide variety of verbs ultimately deriving from the same root, for example $\sqrt{\text{gañ}}$ (*gañ*, *gañ*) ‘fill intr.’, $\sqrt{\text{gañ}}$ (*hgeñs*, *bkañ*, *dgañ*, *khon*) ‘fill tr.’, $\sqrt{\text{keñs}}$ (*kheñs*, *kheñs*) ‘be full’, $\sqrt{\text{sgañ}}$ (*sgañ*, *bsgañs*, *bsgañ*, *sgoñs*) ‘grow, become full’, $\sqrt{\text{skañ}}$ (*skañ*, *bskañs*, *bskañ*, *skañs*) ‘complete, fulfill’. The *s-* prefix is also used to form verbs from nouns (e.g. *nag* ‘speech, talk, word’, $\sqrt{\text{sañag}}$ ‘praise, recommend’).³ Regamey (1946–7: 28, note 4), Bielmeyer (1988: 18), and Nishida (1994: 4) regard the *s-* causative prefix in Tibetan as productive into the historical period; they offer no evidence in favor of this conclusion.

There is some controversy concerning what phenomena in Chinese to compare to Tibetan *s-*. Baxter and Sagart (2012) reconstruct **s-* in Old Chinese for such pairs as 譁 **ŋ⁵⁵ak* ‘speak brusquely’, and 慍 **s-ŋ⁵⁵ak-s* ‘complain, accuse’; 當 **t⁵⁵an* ‘have the value of, rank with’, 商 **s-tan* ‘estimate (v.)’. In contrast, Mei (2012) employs a Chinese **s-* prefix to explain the same pairs for which Sagart reconstructs **N-* (*vide supra*). Thus, where Sagart reconstructs 中 **truŋs* ‘hit the center’ and 仲 **N-truŋ-s* ‘be in the middle’, Mei reconstructs 中 **s-drjəŋgws* and 仲 **drjəŋgws* (cf. Handel 2012: 62, 70).

Phenomena in other languages that have been compared to Tibetan *s-* usually consist of voiced and voiceless pairs of verbs, such as Lolo *bi⁵⁵* ‘come out’, *pi⁵⁵* ‘cause to come out’, Lahu *dɔ³¹* ‘drink’, *tɔ³³* ‘give to drink’, and Burmese *prat* ‘be cut in two’, *phrat* ‘cut in two’ (Mei 2012: 6 et passim). Since Tibetan and Rgyalrong exhibit both voicing alternation and *s-* prefixation (Jacques 2012d: 215), efforts to credit voicing alternation in related languages to a **s-* prefix are unmotivated.

34.14.2 Suffixes

Suffixes are characteristic of nominal derivation. Identifiable nominal suffixes include the segmental suffixes *-s*, *-n*, and *-d*, the suffixes indicating gender *-pa*, *-ma*, *-po*, *-mo*, a diminutive *-bu*, the productive suffixes *-can* and *-ñid*, and the more rarely occurring suffixes *-ka*, *-so*, *-yas*, *-a*, and *-e*; there is also a different suffix *-e* used to form adverbs.

34.14.2.1 Nominal Suffix *-s*

Conrady (1896: 43) points to instances in which *-s* derives nouns from verbs, citing the pairs such as $\sqrt{\text{hgro}}$ ‘go’ and *hgros* ‘gait’, or $\sqrt{\text{lug}}$ ‘pour’ and *lugs* ‘mold, casting.’ Relying particularly on pairs such as *khun* ‘hole’, *khuns* ‘mine’; *khon* ‘inside’, *khons* ‘middle, midst’, Simon (1941: 388) suggests that the suffix *-s* derives from the noun *sa* ‘place’; Beyer repeats this suggestion (1992: 118), but “this etymology does not seem well motivated in most of the cases” (Uebach and Zeisler 2008: 309), and more importantly Simon’s proposal is an error of methodology, confusing diachrony and synchrony. The source of one element of a synchronic system cannot be another element in the same synchronic system. Instead, the source of an element from one stage of a language must be sought in a preceding stage. The *shep-* in ‘shepherd’

³ Emmerick (1983: 68) gives the remarkable counterexample *srubs* ‘a seam’, $\sqrt{\text{hdru}}$ ‘sew’ where the noun occurs with the *s-* prefix, and the verb without.

does not derive from the word ‘sheep,’ but rather from the Old English word *scaep* ‘sheep’ as it developed in compound with *herde* ‘herder.’

Denwood (1986) points to a number of pairs of words to argue for a different -s suffix with a ‘collective meaning’ (e.g. *rnam* ‘individual,’ *rnams* ‘plural marker’; *cha* ‘part,’ *chas* ‘things’). Uebach and Zeisler (2008: 309) identify three uses of -s: (1) a collective suffix, (2) a resultative suffix, (3) a suffix that builds abstract nouns. Although they collect philological attestations of -s suffixed to the second member of compounds, they do not clearly present evidence for each of these three functions. They appear to believe that the use of -s with nominal compounds is either abstract or collective in words such as *rje-blas* ‘official service’⁴ and *yig-tshans* ‘insignia of rank,’ whereas the -s is ‘resultative’ when the second member of the compound is an adjective in such compounds as *khan-bzans* ‘good house’ and *lag-rins* ‘long arms.’⁵ Uebach and Zeisler (2008: 321) connect this “resultative” use of -s with adjectives to the past stem suffix -s of the verbal system. The number of -s suffixes in Tibetan and their functions remains beset with controversy.

Chinese also has an *-s suffix of controversial use. Downer (1959) identifies eight different functions. Schuessler (1985) argues that these proceed from a single category “inversion” comparable to the passive or middle voice in Indo-European languages (e.g. 威 *ʔuj ‘awe-inspiring’ and 畏 *ʔuj-s ‘fear v.’); Mei (2012) argues for the primacy of a perfective meaning of *-s in Chinese as seen in the pair 傳 *Cə-m-tron ‘transmit’ and 傳 *m-tron-s ‘what has been transmitted,’ an analysis that supersedes his earlier proposal for nominalization (e.g. 入 *nəp ‘enter’ and 內 *nʰəp-s ‘inside’) and a change from an “endodirectional” verb (內向動詞) to a “exodirectional” verb (外向動詞) (e.g. 買 *mʰrajʔ ‘buy’ and 賣 *mʰrajʔ-s ‘sell’) as the two primary functions (Mei 1980, cf. Sagart 1999: 131–2). Jacques (2003) compares both Chinese and Tibetan -s to an -s suffix in Rgyalrong, the primary function of which he sees as nominalization.

34.14.2.2 Nominal -n and -d

In some cases a noun is formed from a verb through the addition of an -n suffix (√rgyu ‘move, wander,’ *rgyun* ‘flow, current, stream’; √gci ‘urinate,’ *gcin* ‘urine,’ cf. Beyer 1992: 117).⁶ In other cases a noun is formed from a verb through the addition of a -d suffix (√dro ‘become warm,’ *drod* ‘warmth’; √na ‘be ill,’ *nad* ‘illness,’ cf. Beyer 1992: 117). Lalou (1953) links these two formations by pointing to an alternation between -n and -d in the spelling of single words, for example *chen/ched* ‘big,’ *gsid/gśin* ‘death.’

Some verbal roots participate in several forms of nominal derivation (e.g. √lta ‘look at,’ *ltad-mo* ‘spectacle,’ *ltas* ‘omen’; √rga ‘be old,’ *rgad-po* ‘old man,’ *rgan-mo* ‘old woman,’

⁴ In their survey of previous scholarship on *rje-blas* Uebach and Zeisler (2008) fail to notice Schuessler (1998).

⁵ Walter and Beckwith (2010: 293 note 7) suggest an analysis of *rdo-rins* ‘stele’ as ‘eternal stone’ rather than for ‘long stone’. It is not clear to me how they would interpret *lag-rins* ‘long hands’ or *rkan-rins* ‘long legs.’

⁶ Simon also draws attention to examples where the bases are not verbs but rather locational terms *ya*, *yan* ‘above,’ *pha*, *phan* ‘beyond,’ *tshu*, *tshun* ‘hither’ (1941: 373).

rgas-ka ‘old age’). Such examples hold the most promise to elucidate the semantics of the various types of nominalization.

Sagart (1999: 135) points to a limited number of potential examples of a suffix **-n* in Chinese (e.g. 語 **ŋa?* ‘speak’, 言 **ŋan* ‘speak; speech’). Geilich (1994) discusses further examples of *-n* in Tibetan, Chinese, Lepcha, Lushai, and Tiddim Chin.

34.14.2.3 *The Suffixes -pa, -ma, -po, and -mo*

The suffix *-pa* forms a noun from another noun, meaning ‘associated with N’ (e.g. *rta* ‘horse’, *rta-pa* ‘horseman’, *yi-ge* ‘letter’, *yi-ge-pa* ‘one who holds a letter of office’, cf. Beyer 1992: 117). When suffixed to cardinal numbers this suffix forms ordinals (e.g. *gsum* ‘three’, *gsum-pa* ‘third’; *bcu* ‘ten’, *bcu-pa* ‘tenth’).⁷ This suffix *-pa* can be distinguished from a verbal noun suffix *-pa/-ba*, because the verbal noun suffix undergoes sandhi according to the preceding phoneme (e.g. *byed-pa* ‘doer’ vs. *bya-ba* ‘task’), but the *-pa* denominal noun suffix does not undergo this sandhi. A word like *skyes-pa* ‘human’ is ambiguous; it could be an application of *-pa/-ba* to the verb $\sqrt{\text{skye}}$ ‘be born’ or the suffix *-pa* to the noun *skyes* ‘birth.’ Furthermore, in words such as *rgyal-ba* ‘victor’, the verbal noun suffix *-pa/-ba* has a semantic function parallel to the suffix *-pa*, suggesting that the two may ultimately be of a single origin.

The suffix *-pa* is comparable to the Chinese suffix 夫 **-pa*, of similar function, thus 田 **lʰiŋ* ‘field’ yields 田夫 **lʰiŋ-pa*, which is directly comparable to Tibetan *žin* ‘field’ and *žin-pa* ‘farmer’ (cf. LaPolla 2003: 27); another instance of the suffix in Chinese is the word 魚夫 **ŋa-pa* ‘fisherman’.

In Tibetan the suffix *-pa* coincides with the noun *pha* ‘father’ and the noun suffix *-ma* similarly coincides with the noun *ma* ‘mother.’ Pairs of words such as *nañ-pa* ‘gander’ and *nañ-ma* ‘goose’ suggest that ‘male’ and ‘female’ are the etymological meanings of these two suffixes. The pair *ra* ‘goat’ and *ra-ma* ‘nanny-goat’ shows that *-ma* can indicate the feminine of a species even when the default term does not end in *-pa*. However, many nouns end with *-ma*, which either have no identifiable gender or even normally have a masculine referent (e.g. *rkun-ma* ‘thief’, *bla-ma* ‘guru’).

A rather specialized productive function of *-ma* is its use to mean ‘book associated with a place.’ Thus, the *Zaṅs-glin-ma* was discovered in the *Zaṅs-glin* chapel at Bsam-yaś monastery, and the *Bkaḥ-chems ka khol ma* is a testament (*bkaḥ-chems*) found inside a hole (*khol*) in a pillar (*ka*). This use in book titles and Buddhist coinages such as *bla-ma* ‘guru’ and *mkhaḥ-ḥgro-ma* ‘*ḍākinī*’ suggest that *-ma* was still productive into the Classical period.

The noun suffixes *-po* and *-mo* are also used to distinguish male and female. There are straightforward minimal pairs such as *grogs-po* ‘male friend’, and *grogs-mo* ‘female friend.’ The three words *rgyal-ba* ‘victor’, *rgyal-po* ‘king’, and *rgyal-mo* ‘queen’ show the suffixes *-po* and *-mo* contrasting with the verbal noun suffix *-pa/-ba*. The suffix *-po* is also used in the irregular ordinal number *dañ-po* ‘first’ (cf. *gcig* ‘one’). The pair *kho-bo* ‘I (male speaker)’ and *kho-mo* ‘I (female speaker)’ shows that *-po* at times occurs as

⁷ When suffixed to *gcig* ‘one’ it yields *gcig-pa* ‘unitary, same’ rather than ‘first.’

-bo. This form of the suffix also occurs in *jo-bo* ‘lord’ (cf. *jo-mo* ‘lady’) and *ri-bo* ‘mountain’ (*ri-mo* ‘picture’ is unrelated etymologically). It would be tempting to say that -bo occurs after open syllable roots, but one encounters counterexamples such as *byed-pa-po* ‘agent’.

When suffixed to adjectives the distinction between -po and -mo, rather than indicating gender, distinguishes a concrete and metaphorical use of the adjective, thus *chen-po* means ‘big, large’ and *chen-mo* means ‘great, grand.’ Because authors prefer to be known for writing great books rather than big books, titles such as *Lam-rim chen-mo* ‘Great Stages on the Path (to Enlightenment)’ is a context where *chen-mo* is often met.

According to Beyer words ending in -pa or -pa/-ba tend to form a feminine equivalent with -ma and words with -po to form a feminine equivalent with -mo, whereas nouns ending in “neither will idiosyncratically select one formative or the other, although there appears to be some tendency for native Tibetan words to select -mo and for borrowed expressions and neologisms to select -ma” (Beyer 1992: 125); Beyer does not substantiate this elegant description.

34.14.2.4 *The Suffixes -can and -ñid*

Perhaps the two most productive formative suffixes in Classical Tibetan are -can ‘possessed of’ (Beyer 1992: 121) and -ñid ‘-ness’ (Beyer 1992: 119). Both suffixes occur frequently in Buddhist terminology, for example *sems* ‘mind’ vs. *sems-can* ‘sentient being (lit. possessed of a mind)’ and *ston-pa* ‘empty’ vs. *ston-pa-ñid* ‘emptiness.’

34.14.2.5 *The Diminutive -bu, -hu, etc.*

The noun *bu* ‘son, child’ suffixed to a noun forms a diminutive (e.g. *khron-pa* ‘a well,’ *khron-bu* ‘a small well,’ *bum-pa* ‘water pot,’ *bum-bu* ‘small water pot,’ cf. Uray 1952: 185). The diminutive can add an unpredicted element of meaning (e.g. *rlig-pa* ‘testicles,’ *rlig-bu* ‘scrotum,’ cf. Uray 1952: 185). In some cases the non-diminutive form is unattested or the diminutive derives from a verb stem (e.g. $\sqrt{\text{dril}}$ ‘turn,’ *dril-bu* ‘bell,’ $\sqrt{\text{tor}}$ ‘throw,’ *thor-bu* ‘fragments, miscellanea’). Often the suffix assimilates to the preceding final consonant, *nal-lu* ‘illegitimate child,’ *gžon-nu* ‘young person’ (Uray 1952: 185–6). After open syllables the -b- weakens to -h- [ɣ], and the vowel -a- ablauts to -e- (e.g. *spra* ‘ape,’ *sprehu* ‘monkey,’ *rta* ‘horse,’ *rtehu* ‘colt, mare,’ cf. Uray 1952: 186). Sporadically the [ɣ] subsequently strengthened to, or was respelled as -g- (e.g. *bu-gu* ‘small boy,’ *gri-gu* ‘small knife,’ cf. Uray 1952: 187). Spellings such as *sa-hon* for *sa-bon* ‘seed’ and *šo-ge* for *šo-be* ‘lie, falsehood’ are evidence outside of the suffix -bu for interchange among -b-, -h-, and -g- (cf. Hill 2011b).

34.14.2.6 *Other Nominal Suffixes: -ka, -yas, -so, -a, -e*

There are a number of additional nominal suffixes in Tibetan for which the evidence is more slim. A number of nouns end in -ka, -kha, or -ga (e.g. *ston-ka* ‘autumn,’ *žen-kha* ‘breadth, width,’ cf. Beyer 1992: 133–134) in some cases this morpheme seems analyzable as a derivational suffix (e.g. *than* ‘flat, field,’ *than-ka* ‘painting’). Suffixed to cardinal numbers -ka functions like ‘-some’ in English (e.g. *gñis-ka* ‘both, twosome,’ *gsum-ka*

'threesome'). Insufficiently distinguishing diachrony and synchrony, Beyer (1992: 133) proposes to derive this suffix from the word *kha* 'mouth.'

Beyer (1992: 122) offers several examples of the suffix *-yas* '-less': *bgrañ* 'number,' *bgrañ-yas* 'numberless,' *mthah* 'end, border, limit,' *mthah-yas* 'endless,' *dpag* 'measure,' *dpag-yas* 'measureless,' *gžal* 'weight, measure,' *gžal-yas* 'infinite.' The name of Tibet's first monastery, Bsam-yas 'inconceivable,' is another instance of the suffix.

Simon (1941: 386) points to a suffix *-so* that appears to derive nouns from nouns or verbs in a very limited number of examples (e.g. *ñan-so* 'hell,' *hkhurul-so* 'mistake' vs. *ñan* 'bad,' *√krul* 'be mistaken').

It is necessary to recognize a suffix *-a* in order to account for certain lexical doublets, *bug*, *bu-ga* 'hole,' *chog*, *cho-ga* 'ritual' (cf. Simon 1941: 386, LaPolla 2003: 25). In examples such as *dar-ba*, *dar-ra* 'buttermilk' and *žal-ba*, *žal-la* 'clay' it seems there is not in fact a suffix *-a*, but rather an assimilation of the suffix *-ba* to the final consonant of the preceding syllable. However, an example like *√htag* 'weave,' *tha-ga-pa* 'weaver' would be difficult to analyze without such a suffix. Similar reasoning requires a suffix *-e* in the analysis of the pair *yig*, *yi-ge* 'letter.'

34.14.2.7 The Adjectival and Adverbial Suffix *-e*

Uray (1953: 232–3) identifies a *-e* suffix that forms adjectives and adverbs from verbal and nominal roots; examples include *khriḡs-se* 'plentiful' from the verb *√kriḡ* 'cohere,' *log-ge* 'groundless, false' from the verb *√log* 'reverse,' and *hbol-le* 'easily, quietly' from *hbol-po* 'soft, smooth'. Other nominal suffixes may be applied subsequent to *-e*, yielding words such as *sdug-ge-ba* 'the state of being somewhat pleasing' from *sdug-pa* 'attractive, agreeable' and *che-ge-mo* 'such a one' from *gcig* 'one' (Uray 1953: 235–6). This adjectival and adverbial *-e* suffix is more characteristic of the spoken dialects and Classical Tibetan influenced by spoken forms; it is seldom met in Old Tibetan (cf. Uray 1953: 240).

34.14.3 Circumfixes and Infixes

Tibetan has only one circumfix of rather restricted employment. A collective meaning 'children of' is applied to certain kinship terms by prefixing an *s-* and suffixing a *-d*: *pha* 'father,' *pha-spad* 'father and children'; *ma* 'mother,' *ma-smad* 'mother and children'; *khu* 'uncle,' *skud-po* 'brother-in-law.' Although they do not quite follow the pattern the three words *phu* 'elder brother,' *spun* 'sibling,' and *span-spun* 'brother's relatives' also merit mentioning in this context (cf. Nagano 1994: 110–11).

Gong proposes a palatal infix with honorific meaning ([1977] 2002), the results of which are most conspicuous after velars, *skam* 'dry,' *skyem* 'be thirsty' (hon.). In other contexts the infix *-y-* is only clear if one bears in mind certain changes in Tibetan historical phonology, for example *dugs* 'light, kindle,' *žugs* < **dyugs* 'fire' (hon.) (cf. Gong [1977] 2002, Hill 2011a).

34.15 DUBIOUS INHERITED PREFIXES

Previous authorities discuss a number of ‘animal’ and ‘body part’ prefixes. Conrady (1896: 40) compares Tibetan *s-* to a *sa-* prefix in Lepcha, which occurs in several animal names, for example *sahu* ‘monkey,’ and *saryóm* ‘otter,’ to which he compares Tibetan *sram* ‘otter.’⁸ Benedict (1972: 106) believes that “with words for parts of the body and animals” *s-* derives from “**sya* ‘flesh; animal’” (1972: 106), his reconstruction for the ancestor of Tibetan *śa* ‘meat, deer.’ For body parts Benedict points to *skra* ‘hair,’ *sñiñ* ‘heart,’ *sna* ‘nose,’ and *so* ‘tooth,’ and for animals *sram* ‘otter’ and *śig* ‘louse’ (1972: 106–8, also cf. Beyer 1992: 115). Chang Kun (1972: 336) rejects Benedict’s prefix as “a dubious proposition”. Languages other than Tibetan present more direct evidence for a sibilant ‘flesh’ prefix. The Burmish language Atsi has a full lexeme *syo*¹¹ ‘flesh,’ which reduces to *si*¹ in compounds (e.g. *si*¹-*cu*³¹ ‘fat meat,’ *si*¹-*mau*³⁵ ‘body hair, fur,’ *si*¹-*mi*¹¹ ‘tail,’ cf. Lustig 2010, vol. II: 335–7, 354). Baxter and Sagart (2012: 45) find an exponent of a putative **sya* ‘flesh’ prefix in Chinese with the pair 節 **ts*⁵ik ‘joint of bamboo’ and 膝 **s*-*ts*ik ‘knee.’

An *s-* is not the only candidate for a ‘body part’ prefix in Tibetan. Laufer (1914: 109) suggests that the “prefix *m-* appears in connection with words denoting organs of the body,” giving examples such as *mchin-pa* ‘liver,’ *mche-ba* ‘tusk,’ and *mcher-pa* ‘spleen.’ Shafer (1938: 17) identifies this prefix with the word *mi* ‘person,’ despite Wolfenden’s (1928: 278–9, cf. Beyer 1992: 114) previous rejection of this association. Shafer (1938: 13) retains *mche-ba* ‘tusk’ as an example despite the infrequency with which humans bear tusks; Shafer and Beyer include *mjug* ‘hind, tail,’ which Beyer redefines as ‘buttocks’ to appear more human. Taking another liberty, Shafer and Beyer give *mtshan* ‘genitals’ as a further example of an *m-* ‘human body’ prefix, even though this meaning is a specialized use of the word ‘sign, signifier,’ occurring in the compounds *mo-mtshan* ‘sign of the female’ and *pho-mtshan* ‘male signifier.’ Wolfenden is probably correct that the “*m-* forms show far too widely sundered senses for the prefix to have ever functioned as an index of mental association” (1928: 279).

Jacques proposes the alternative ‘animal’ and ‘body part’ prefixes *g-* and *d-* respectively (2008: 53–4). In support of the *g-* “animal” prefix he compares Japhug Rgyalrong *qa-ljaḅ* ‘eagle,’ *qa-zo*, ‘sheep,’ and *qro* ‘ant’ with Tibetan *glag*, *g.yaṅ*, and *grog-mo* of the same meanings. In the case of the body part prefix *d-* Jacques does not present direct Tibetan cognates but cites the Japhug words *tu-mtshi* ‘liver’ and *tu-mkhe* ‘neck’ in support for analysing a *d-* prefix in the Tibetan words *gdoñ* ‘face,’ *dbu* ‘head,’ and *dpuñ* ‘back’ (2008: 57).

It is unlikely that Tibetan has three body part prefixes *s-*, *m-*, and *d-*, and two animal prefixes *s-* and *g-*. None of these proposed prefixes is productive in Tibetan, and they do

⁸ The Lepcha prefix occurs in several animal names (*saka* ‘deer,’ *sacák* ‘leopard,’ etc.), but the only potential example in a body part is *sagrek* ‘throat, intestine’ (Plaisier 2007: 239).

not form body parts and animals from other words, so are not derivational *per se*. Such proposals require more methodological reflection. A similar approach to English would likely yield an *h*- ‘human body part’ prefix in the English words ‘hand,’ ‘hair,’ ‘head,’ ‘heart,’ ‘hip,’ and ‘heel.’ Similarly, an additional Tibetan *r*- ‘animal’ and ‘body part’ prefix could be proposed on the basis of *rkañ* ‘leg,’ *rna* ‘nose,’ and *rkyañ* ‘wild ass.’

34.16 GHOST AFFIXES

Wolfenden (1928: 277, cf. 1929: 30) points to the verb *mñah* ‘to have,’ *mkhyen* ‘know’ and *mgu* ‘be glad’ as examples of an “inactive intransitive” *m*- prefix. Matisoff (2003: 89) similarly sees an opposition “between inner-directed or stative verbs” with *m*- and “transitive or outer directed or causative verbs” with *s*-. Sagart (2006) attempts to link Tibetan *m*- with Chinese *N-, a prefix that forms intransitive verbs out of transitive verbs that he also links to Tibetan voicing alternation (see Section 34.14.1).⁹ However, it is a mistake to associate *m*- with intransitives; Snellgrove (1954: 199) draws attention to such transitive examples as *mthoñ* ‘see,’ *mchod* ‘honor,’ and *mdzad* ‘do’ (cf. Hill 2009: 175). If *m*- has any morphological function in Tibetan it remains obscure.

Many authors posit a *-d* suffix causative (Conrady 1896: 45, Laufer 1914: 62, Róna-Tas 1985: 178, Beyer 1992: 176, LaPolla 2003: 24, Matisoff 2003: 457–88); this is a mistake which results from confusing the *-d* present tense suffix in some transitive paradigms with derivation (Hill 2009: 176). Thus, in the pair of *hgye*, *gyes*, *hgye*, ‘be dispersed’ and *hgyed*, *bkyes*, *dgye*, *khyes* ‘disperse’ the *-d* suffix is an inflectional suffix and not a causative formative. LaPolla (2012: 12) posits a **-t* transitivizing suffix as a ‘distinctive feature’ of Sino-Tibetan; whether or not pairs such as Rawang *ṅuī* ‘weep’ and *ṅut* ‘mourn, cry for’ (LaPolla 2003: 24), Kachin *māni* ‘laugh’ and *mānit* ‘laugh at’ (Benedict 1972: 101), or Limbu *-sṅṅ-* ‘be done (food)’ and *-sṅṅd-* ‘cause to become done (food)’ (van Driem 1987: 249–67) warrant the reconstruction of such an affix, Tibetan does not support it.

PART 3 RGYALRONG

JACKSON T.-S. SUN

Rgyalrong is a morphologically complex language spoken by well over 100,000 ethnic Tibetans residing in Aba and Ganzi prefectures in northwestern Sichuan. The

⁹ Sagart (1999: 76–86) also reconstructs a **m*- volitional prefix in Chinese, but appears never to have proposed linking this with Tibetan *m*-.

language is internally diverse, comprising at least four distinct language-like “dialects”: Situ (Eastern), Japhug (Northeastern), Tshobdun (Northwestern), and Showu (Western). It is in turn closely affiliated with two neighboring languages Horpa and Lavrung to form a Rgyalrongic unit under the Qiangic branch in Sino-Tibetan (Sun 2000a, 2000b).¹⁰

Rgyalrong word structure is agglutinative and synthetic, with some degree of fusion. Nominal inflectional categories include number, case, and head-marked possession. Dependent case marking is under-developed and subject to ellipsis where ambiguity does not arise. Most verbs are rigidly sub-classified as either transitive or intransitive.¹¹ Verbal inflectional categories include person-number, polarity, direction (direct vs. inverse), topography-based spatial deixis, transitivity, tense-aspect, and evidentiality. Person marking follows an empathy-based hierarchy (DeLancey 1981, Sun and Shi 2002). The syntax features head-marking tendencies and a head-final word order with pragmatically determined variations. Nouns, verbs, ideophones, and particles constitute the major word-classes of the language. Property words (equivalent to adjectives in other languages) are a sub-class of stative verbs. Nominalization occupies a central role in the architecture of the grammar, as elsewhere in Sino-Tibetan (Matisoff 1972, Noonan 1997, Genetti et al. 2008).

A majority of lexemes are internally complex, formed by compounding, noun incorporation (amalgamation of a nominal stem into a verb stem; see Jacques 2012c), as well as derivation—the morphological process that creates new lexical meanings, alters word-classes, or produces sub-classes within a word-class. It will be the aim of this chapter to overview the last of these word-formation techniques in Rgyalrong.

Stems, the forms of a lexeme to which derivational processes apply, undergo extensive morphological alternation, the topic of Section 34.17. Section 34.18 is devoted to describing the rich variety of formal devices utilized in derivation. This is followed by a detailed look at the two major functional types into which derivation falls (Section 34.19). Ideophones deserve special treatment, as they constitute a distinct minor word-class in the language, characterized by unique patterns of morphological variation and trans-categorial operation (Section 34.20). Conclusions are drawn in Section 34.21.

¹⁰ Unless otherwise noted, the Rgyalrong data cited in this chapter are in the *Qbc⁴weri²* variety of the Tshobdun dialect, from field notes collected during my extended fieldwork in Sichuan. Tshobdun is spoken by around 3,000 agriculturalist Tibetans residing in seven villages within Tshobdun (in Chinese Caodeng 草登) Township. Transcription of the data is phonemic, using symbols in their usual IPA values. The default (pitch) accent position in Tshobdun is stem-final; marked penultimate accent is indicated by an acute accent ´; there is also a functionally restricted falling tone, marked by ˘. The two tonemes in Showu are level ˉ and falling ˘. The interlinear morpheme glosses follow the Leipzig Glossing Rules, except for the following: CONT: continuous aspect; DM: discourse marker; GP: generic person; HTR: high transitivity; INV: inverse; LTR: low transitivity; MED: meditative evidential; SPON: spontaneous; SUB: subordinator.

¹¹ For useful discussions of transitivity lability in the Japhug dialect, see Jacques (2012b: §7).

34.17 STEM FORMATION

34.17.1 Nominal Stems

Many (but not all) native Rgyalrong nouns take the form of a nominal stem prefixed with an absolutive nominal prefix *tə-*.¹² Only bare noun stems occur in inalienably possessed or derived forms, or non-initially inside compounds, as shown by the examples in (1) containing the nominal root *rjev* ‘wife’:

- (1) *tə-rjev* ‘wife (noun in absolutive form)’
ɐ-rjev ‘my wife (inalienably possessed noun)’
mɐ-rjev ‘marry into one’s husband’s household (denominal verb)’
kəndʒə-rjə-rjev ‘married couple (derived noun denoting binary human relations)’

The synchronically opaque formatives *qɐ* and *kə-*¹³ co-occur with certain stems denoting objects, animals, and plants:

- (2) *qɐⁿbət* ‘sand’ *kəxʃiʔ* ‘basket carried on back’
qɐperʔ ‘dhole’ *kəʃpet* ‘marmot’
qɐʃ^he ‘peach’ *kəʃnəmʔ* ‘ear (top part of grain plant)’

These elements could once have been nominal prefixes in their own right, comprising a system of nominal classification with *tə-*, but have now fused into the noun stems, as evidenced by their retention in derived forms like *nə-qɐʃwɛʔ* ‘catch fish,’ from *qɐʃwɛʔ* ‘fish.’¹⁴

34.17.2 Verbal Stems

Verb-stem differentiation is a highly visible feature of the language, exhibiting a level of richness rarely encountered elsewhere in Sino-Tibetan. Many common verbs in the more conservative forms of the language show functionally distinct verb stems. Comparative evidence unveils three such stems across the dialects, referred to here as Stem 1 (verb base), Stem 2 (past stem), and Stem 3 (singular transitive non-past stem).¹⁵ Stem 2 is functionally most pervasive, occurring in the perfective, imperfective past,

¹² The term *absolutive* is borrowed from Uto-Aztecan linguistics. This and a few other prefixes in Rgyalrong (including the causativizer *sə-*, and verbalizers *nə-*, *wə-*, and *rə-*; see further sections) all exhibit a variant with *-ɐ-* vocalism, stemming from fusion of the prefix vowel *-ə-* and a stem-initial *ɐ-*. The formative *-ɐ* could once have served some stem-building function but has become in large measure lexically determined (e.g. *tə-mu* ‘rain’ vs. *tɐ-worʔ* ‘hoarfrost’; *tə-xtse* ‘shoes’ vs. *tɐ-xtɐ* ‘hat’; *tə-mnu* ‘awl’ vs. *tɐ-qɐv* ‘needle’).

¹³ Cf. *k^hə-* in the Situ dialect.

¹⁴ Comparative data indicate that the true ‘fish’ root is *ʃwɛʔ*, cf. Situ *-wjõ*, Yelong Lavrung *-jo*.

¹⁵ The non-base (second and third) stems are indicated with subscript numerals.

subtypes of continuous-aspect verb forms and, depending on the dialect, also in converbs (nonfinite verbal forms serving an adverbial function) and oblique deverbal nouns, while Stem 3 occurs only in certain highly specific transitive contexts constrained by number (singular), tense-aspect (non-past, non-progressive), as well as direction (direct). Stem 1 occurs in all other contexts not selected by the marked stems. Verb-stem alternation is best illustrated from the conservative Showu dialect, which preserves the original complexity to a high degree. As observed below, the marked stems in this dialect are formed through a number of stem-modification processes involving vocalic, consonantal, and prosodic adjustments, as well as subtraction and suffixation:

(3) Showu verb-stem alternation

Stem 1	Stem 2	Stem 3	
<i>ts^hə</i>	<i>ts^hɿ-t</i>	---	'say'
<i>smô</i>	<i>smə-t</i>	---	'stay; stop'
<i>rkwə̃t</i>	<i>rk^hũ</i>	<i>rkô</i>	'put into (a container)'
<i>rbā</i>	<i>rbî</i>	<i>rbē</i>	'press'
<i>fû</i>	<i>fō</i>	<i>fə-m</i>	'entrust'
<i>ɤartî</i>	<i>ɤart^hɿ</i>	<i>ɤarte</i>	'wear (a hat)'

Verb-stem derivation in the Tshobdun dialect is less diverse but more thoroughgoing. With a few exceptional accented verbs with invariant stems,¹⁶ Stem 2 is formed principally by regular *glottality inversion*, whereas Stem 3 is identical to Stem 1 except for a different vowel grade:

(4) Tshobdun verb-stem alternation

Stem 1	Stem 2	Stem 3	
<i>χtu</i>	<i>χtuʔ</i>	<i>χtɛ</i>	'buy'
<i>fɕɛʔ</i>	<i>fɕɐ</i>	<i>fɕɐʔ</i>	'change'
<i>rɐ'ski</i>	<i>rɐski</i>	<i>rɐ'skɛ</i>	'pull'

Nominalizations of verbs are built on either Stem 1 or Stem 2 (see Section 34.19.2.1). However, a number of common verbs show an additional irregular stem in deverbal nouns marked by *tə-*, with a shifted vocalism and a suffix *-t*.¹⁷

(5) Tshobdun irregular verb stems

Stem 1	Stem 2	Deverbal noun	
<i>mərku</i>	<i>mərkuʔ</i>	<i>tə-mərka-t</i>	'steal'
<i>ɛⁿdu</i>	<i>ɛⁿduʔ</i>	<i>tɐ-ɛⁿdə-t</i>	'punch'
<i>rŋi</i>	<i>rŋiʔ</i>	<i>tə-rŋə-t</i>	'borrow'

¹⁶ For example: *mts^hỗt* 'be full'; *tə-mts^hỗt* 'have become full [PFV]'.
¹⁷ See 34.18.2.2.2 below.

34.18 MORPHOLOGICAL DEVICES OF DERIVATION

34.18.1 Affixation

Rgyalrong has a strong preference for prefixes over suffixes in all areas of morphology. Suffixation occurs only in deriving certain sub-classes of nouns (e.g. *pyɛ-pu* ‘chick’ with the diminutive *-pu*; cf. *pye* ‘bird’), and as an irregular stem-building device in verb-stem formation (see Section 34.17.2). The highly synthetic nature of Rgyalrong morphology is illustrated by words containing long strings of prefixes, such as:

- (6) *tʰv-kə-o-nə-ɣv-sv-ɞʰduʔ* *tɛlɲaʔ=nəʔ*
 CONT:LTR-NMLZ:SBJ-INV-SPON-REFL-CAUS-beat₂ child=DET
 ‘the child who is getting himself beaten’

The relativized verb in example (6) shows a stack of three derivational prefixes (a reflexive, a spontaneous, and a causative) and three inflectional prefixes positioned peripherally to the derived stem.

34.18.2 Morphological Processes

34.18.2.1 Reduplication

Reduplication, or iteration of some part of the phonological material of a root, is a prevalent morphological process in Rgyalrong.

Full reduplication of the root is drawn upon heavily to form ideophones (Section 34.20), and verbal iterative aspect (e.g. *nɐ-spɛr-spɛr* ‘keep moving (objects)’ with prefix *nɐ-*, root *spɛr* ‘move (an object)’). It also occurs in miscellaneous lexicalized derivations, such as *mer-ko-mer* ‘every night’ (root *mer* ‘night’) and *snj-ku-snj* ‘everyday’ (root *snj* ‘day’).

Much more extensively attested are partially reduplicated syllables of an alliterative type comprising the onset of the root (or final syllable of a disyllabic root) plus a vowel *-ə*,¹⁸ as seen in the intensive (e.g. *ʰbrə-ʰbriʔ* ‘be very tall’; root *ʰbri* ‘be tall’) and comparative (e.g. *ʰbrɛ-ʰbriʔ* ‘be taller; more tall than average’) stative verb forms. This semantically multi-valued reduplication type often operates in unison with other morphological

¹⁸ Or *v-* in the comparative formation seen immediately below. Some verbs manifest an irregular reduplicative vowel *-ə*, as in the ambulative form of the verb *prɛt* ‘break (sth pliable)’; *nɐ-prɛprɛt* ‘move about breaking (pliable things)’; cf. the comparable irregular reduplicative vowel *-oɣ* reported in Japhug (Jacques and Chen 2004: 9).

devices, for instance with ablaut in deriving nouns meaning “various kinds of N” (e.g. *pyəpye* ‘various kinds of birds’ from root *pye* ‘bird’). Partial reduplication accompanied by affixation operates productively to form a wide range of derived nouns and verbs. Examples include terms of binary human relations (e.g. *kəⁿdʒə-tʃətʃu* ‘father and son’ with complex prefix *kəⁿdʒə-*; root *tʃu* ‘son’), denominal comitative adverbials (e.g. *ɣgerɣen ko-slomame* ‘the master together with his disciple’ with prefix *ko-*; root *slome* ‘disciple’), converbs (e.g. *brɛzʔ kə-rkərko kʰi* ‘hand over (sth) with rice contained in it’ with prefix *ko-*, root *rko* ‘put into a container’; *sɛ-nɛrərəʔ nəʃev* ‘fall asleep laughing’ with prefix *sɛ-*, Stem 2 of root *rɛ* ‘laughter’), as well as pluractional (e.g. *sɛ-ⁿdʒə-ⁿdze* ‘eat various things’ with prefix *sɛ-*, root *ⁿdze* ‘eat’), ambulative (e.g. *nɛ-ɁməɁmɛrʔ* ‘move about smearing (sth)’ with prefix *nɛ-*, root *Ɂmɛrʔ* ‘smear’), concessive (e.g. *ʃo nɛ-nə-ŋəŋo* ‘whoever (he/she) is’ with prefix *nə-*, root *ŋoʔ* ‘be’), and crescendo (e.g. *nə-tʃətʃot* ‘be hotter and hotter’ with prefix *nə-*, root *tʃot* ‘be hot’) verb forms.

Partial reduplication of a rhyming type is also attested, where the verb root takes the prefix *nɛ-* and a following reduplicated syllable with an *-l* onset and a rhyme copied from the root. This structure applies to motion verbs to impart a distinct reciprocating (i.e. back and forth) meaning:¹⁹

(7)	<i>stʰɔ</i> ‘push’	<i>nɛ-stʰɔ-lɔ</i> ‘push back and forth’
	<i>nə-mder</i> ‘jump’	<i>nɛ-mdɛr-ler</i> ‘jump back and forth’
	<i>ⁿdʒev</i> ‘roll’	<i>nɛ-ⁿdʒev-lev</i> ‘roll back and forth’

34.18.2.2 Apophony

Apophony refers to phonological alternations within a word. In Rgyalrong, vowels, consonants, as well as prosodic features undergo apophonic changes to produce morphologically related forms.

34.18.2.2.1 Vocalic Apophony

As exemplified in Section 34.17.2 above, vocalic apophony (ablaut) is used extensively in forming verbal and nominal stems. Rgyalrong dialects vary considerably in the number of observed verbal ablaut series, ranging from none in the Situ dialect, four or five in the Japhug dialect (Lin and Luo 2003: 26; Jacques 2008: 228–35), seven in the Tshobdun dialect, to almost fifty in the Showu dialect (Sun 2004b). Ablaut also supplements reduplication to produce derived words meaning ‘various kinds of’:

(8)	<i>pye</i> ‘bird’	<i>pyəpye</i> ‘various kinds of birds’
	<i>sekoʔ</i> ‘bush’	<i>sekəkoʔ</i> ‘various kinds of bushes’
	<i>vdɛ</i> ‘be well’	<i>vdɔvdɛ</i> ‘be well in various ways’

¹⁹ Cf. Jacques and Chen (2004: 9).

34.18.2.2.2 *Consonantal Apophony*

Consonantal alternation occurs most frequently in verb-stem derivation. The Tshobdun formation of Stem 2 via glottality inversion offers a striking illustration of productive consonantal apophony:

(9) Tshobdun regular glottality inversion

<i>Stem 1</i>	<i>Stem 2</i>	
<i>pe</i>	<i>peʔ</i>	‘do’
<i>sroŋ</i>	<i>sroŋʔ</i>	‘guard’
<i>toʔ</i>	<i>to</i>	‘exist’
<i>pʰjəzʔ</i>	<i>pʰjəz</i>	‘wipe’

Other prime examples of consonantal apophony are from the Showu dialect, including aspiration of stop onsets, excrescent coda *-t* in Stem 2, as well as excrescent coda *-m* in Stem 3, see example (3) above. Another notable type—vestigial remains of an ancient Sino-Tibetan distinction (LaPolla 2003: 1.2)—is represented by a few lexical causative pairs marked by onset voicing alternation, where the non-causatives contain voiced pre-nasalized onsets while the corresponding causatives show plain or aspirated voiceless onsets:

(10) Tshobdun onset voicing apophony

<i>Non-causative</i>	<i>Causative</i>	
<i>ᵿbrət</i>	<i>prət</i>	‘break (a rope)’
<i>ᵿovʔ</i>	<i>cʰovʔ</i>	‘break (a stick)’
<i>ᵿbəm</i>	<i>pəm</i>	‘overflow’
<i>ᵿba</i>	<i>pʰa</i>	‘burst’

34.18.2.2.3 *Apophony in Prosody*

Rgyalrong dialects vary considerably in their word-level prosody types. Although pitch distinctions serve only a minor function in the Japhug dialect (Jacques 2008: 36–7), full-fledged tonal (in Situ and Showu; Lin 2012; Sun 2004b: 272) and pitch accent (in Tshobdun; Sun 2003: 491) systems are attested elsewhere in Rgyalrong.

As shown in example (3), Showu makes use of pervasive but irregular tonal alternations, among other phonological devices, to derive Stem 2 and Stem 3 from the verb base. In Tshobdun, relocation of (default) final to penult accent regularly derives measure words from container nouns:

(11) Tshobdun accent shift in derived measure words

<i>Container noun</i>	<i>Measure word</i>
<i>pʰərtfu</i> ‘bowl’	<i>pʰərtfu</i> ‘bowlful’
<i>tʂoŋtʂoŋ</i> ‘cup’	<i>tʂoŋtʂoŋ</i> ‘cupful’

Accent shift also occurs sporadically in lexical derivation, as in word pairs like *nənəʔ* ‘that’ versus *nənə* ‘there,’ and *kə-tforʔ* ‘be sour’ versus *kə-tfor* ‘dried pickles.’

34.18.2.3 Suppletion

Our overview of formal devices utilized in Rgyalrong derivational morphology will not be complete without mentioning suppletive, or phonologically unrelated, forms in verb-stem alternation, as suppletive stems may be required in verbal derivation of the category-preserving as well as category-changing types. Several important characteristics of verbs with suppletive stems are noted. First, they are all intransitive, including certain existential verbs and a few basic motion verbs. Second, the observed cases of suppletion belong to both the “categorical” (where suppletive stems are tied to a single grammatical category) and the “non-categorical” types (Veselinova 2007). The positive existential verb *wɛʃo* ‘be seen/be found’ and its negative counterpart *məŋɐ* ‘be not seen; disappear’ illustrate the former type. Suppletive stem alternation of the non-categorical type can be observed in the following Showu data, where the alternation of irregular stems involves not only typologically common tense-aspect categories, but also a unique feature of *orientation specificity*, by which orientationally explicit events (i.e. ‘go in a specific direction’) are opposed to orientationally implicit ones (i.e. ‘go away’). Importantly, the distinct stems are not distributed in a straightforward “one stem per category” fashion:

(12) Showu suppletive stems

<i>Stem 1</i>	<i>Stem 2</i>	<i>Progressive</i>	
<i>xwɛ̃</i>	<i>xwɛ̃t</i>	<i>xwɛ̃t</i>	‘go (away)’
<i>rɪ̃</i>	<i>rɛ̃t</i>	<i>xwɛ̃</i>	‘go (in a specific direction)’
<i>vɛ̃</i>	<i>vɪ̃</i>	<i>vɛ̃</i>	‘come (over)’
<i>tɪ̃</i>	<i>vɪ̃</i>	<i>tɔ̃</i>	‘come (in a specific direction)’

34.19 FUNCTIONAL TYPES OF DERIVATION

Derivation falls under two broad types. The category-preserving type creates derivatives belonging to sub-classes of the same word-class, whereas the category-changing type converts one word-class to another (Bybee 1985: 82; Aikhenvald 2007: 40–4).

34.19.1 Category-preserving Derivation

Derivation of this type applies in some measure to nominal roots. Several sub-classes of nouns are created by means of derivational affixes, including gender (male *-pʰo*, female *-mo*) and diminutive (*-pu*; *-tse*; *-lo*) suffixes, and the complex prefix *kɐ̃ndʒə-* marking

terms of binary human relations seen in Section 34.18.2.1. Nominal number (dual and plural) marking in Rgyalrong is better considered a derivational rather than an inflectional process, owing to its grammatical optionality²⁰ and its *associative plural* meaning (e.g. *stenlo=rɛʔ* ‘Stanlo and associates’).

Morphology that generates sub-classes of verbs is much more pervasive. Of particular importance to the grammar are derivations that alter verb valency (Sections 34.19.1.1–34.19.1.2) or produce stative verbs expressing propensity (Section 34.19.1.3).

34.19.1.1 Valency-increasing Derivation

Rgyalrong contains a set of prefixes that function to introduce a core participant into the argument structure of the input verb.

34.19.1.1.1 Valency-increasing Prefix *nə*-²¹

Among its miscellaneous functions, the prefix *nə*- creates a derived transitive verb from an intransitive verb-stem. Derived verbs of this kind may contain introduced *animate* subject arguments to impersonal verbs (e.g. *nɛ-rtaʔ* ‘regard as sufficient’, from root *rtaʔ* ‘be sufficient’); in another important function, the structure works as an *applicative* construction where a peripheral locative argument is promoted to a core-object argument:

- (13) a. *k^hezeʔ* *ɛ̃-ta* *t^hɛ-vⁿdzət*
 dog 1SG:POSS-top CONT:LTR-bark₂
 ‘The dog is barking at me.’
 b. *k^hezeʔ=kə* *t^hɛ-o-nɛvⁿdzət-aŋ*
 dog=ERG CONT:LTR-INV-bark.at₂-1SG
 ‘The dog is barking-at me.’

In (13b), the locational argument ‘I’ of the intransitive verb *ɛ̃ⁿdzət* ‘bark’ in (13a) is coded as an applied object to the derived transitive verb *nɛ-vⁿdzət* ‘bark-at’, triggering inverse and object person marking on the verb, as well as ergative-case marking on the transitive subject ‘dog.’

34.19.1.1.2 Causatives

Barring a handful of lexical causative pairs (Section 34.18.2.2.2), causativization is achieved morphologically by means of prefixes that add a subject argument (the causer) to the verb. Three functionally distinct causative prefixes are found across the dialects: *sə*-,²² *ʃə*-,²³ and *wɛ*-. While *sə*- appears to be the default causativizer applicable to

²⁰ Unlike in Indo-European languages, nominal number marking in Rgyalrong is not directly linked to verb agreement, and generally does not occur with inanimate or quantified nouns.

²¹ Homophonous with the general verbalizing prefix discussed in Section 34.19.2.2.1.

²² When attached to roots with a non-cluster onset, the causative prefix *sə*- in Tshobdun shows an added coda -y.

²³ This usage is first noted in the Somang variety of Situ (Jin et al. 1958: 83). Reflexes of this prefix are also reported in Japhug, but apparently without the ‘direct involvement’ meaning (Jacques 2008: 72–3).

all verb stems, the other prefixes are restricted in function. The prefix *ʃə-* (and its positional variants) implies direct involvement of the causer (e.g. *ʃ-pʰi* ‘physically assist sb to escape’, cf. *səy-pʰi* ‘allow or cause sb to escape’). The prefix *wə-*, on the other hand, causativizes inanimate intransitive verbs only and may contrast with *sə-* in meaning:

- (14) a. *cʰɛʃi ne-kə-səy-cʰiʔ=nəʔ mimʔ=cə*
 beer IMPV-GP-CAUS-be.sweet=SUB be.tasty=MED
 ‘Beer is tasty when one allows it to sweeten (naturally and gradually).’
 b. *cʰɛʃi ne-kə-wə-cʰiʔ=nəʔ mimʔ=cə*
 beer IMPV-GP-CAUS-be.sweet=SUB be.tasty=MED
 ‘Beer is tasty when one sweetens it (e.g. by adding sugar).’

Here, the general causativizing prefix *sə-* denotes causation of an increase in the degree of the predicated state (“make sth sweeter”), whereas *wə-* simply indicates causation of a changed state (“make sth sweet”).

34.19.1.2 Valency-decreasing Derivations

Rgyalrong employs several derivational patterns that downplay the core arguments of a transitive verb, turning it into a derived intransitive.

34.19.1.2.1 Agentless Passive

The agentive argument of a transitive verb may be removed by a derivational process in Rgyalrong, resulting in a passive-like construction. This agentless passive is highly constrained in form, comprising Stem 1 plus a dedicated marker *ɐ-* and optional evidential marking.²⁴ The construction conveys a specific *resultative* meaning—to speculate on a past action performed by unknown agents based on tangible results of that action. As an illustration, consider the following utterance made by someone walking into the room to find empty alcohol bottles on the table:²⁵

- (15) *kʰəkʰoz tɛwaʔ kəlden ɐ-tʰi=cə jɐ*
 here alcohol much PASS-drink=MED DM
 ‘A lot of alcohol was consumed here!’

34.19.1.2.2 Antipassives

Valency reduction can also be achieved by antipassive structures which eliminate the object argument from the clause. Antipassives in Rgyalrong belong to the semantic/pragmatic type (Cooreman 1994), as their main function is to downplay indefinite or generic objects in order to focus on the activity itself (Dixon 2010: 168). Two

²⁴ The Japhug passive verb can be specified for perfectivity (Jacques 2012a: §4). This is not possible in the Tshobdun dialect.

²⁵ For agentless passive examples in Japhug, see Jacques 2008: 76–7, 2012a: §4.

antipassive prefixes *sɐ-* and *rɐ-* are in contrast, referring respectively to a [+HUMAN] and a [-HUMAN] demoted object. They turn a transitive action verb into an intransitive action verb:

- (16) *ntʰe* 'kill'
 sɐ-ntʰe 'do killing (of people)'
 rɐ-ntʰe 'do slaughtering (of livestock)'

34.19.1.2.3 *Reflexive and Reciprocal Formation*

Reflexive and reciprocal constructions present marked scenarios in terms of transitivity, as the subject and object arguments are identical (reflexive) or interchangeable (reciprocal). Both structures are produced in Rgyalrong by morphological operations, resulting in derived verb stems with reduced transitivity.

The reflexive verb is marked by a prefix complex *oʃɐ-*, further analyzable into the inverse *o-* plus a reflexive prefix *ʃɐ-*.²⁶ Formation of the reciprocal verb also involves a prefix complex *ɐ(mə)-*.²⁷ attached to a reduplicated root.²⁸ In the following examples, reduction of transitivity in these derivations is evident from the absence of pronominal objects, and the choice of the low-transitivity prefix *tʰɐ-* when the verb is in the continuous aspect:

- (17) a. *oʃiʔ* *tʰɐ-o-ʃɐ-ɐʰdu=cə*
 3SG CONT:LTR-INV-REFL-punch₂=MED
 'She/he is punching her/himself.'
 b. *kəxtʃipu=niʔ* *kʰɐxtə-ʰdi* *tʰɐ-ɐ(mə)-stʰəstʰɔʔ-ʰdzə* *ʃɐ*
 small.child=DU roof.top.platform-edge CONT:LTR-RECP-push₂-3DUDM
 'Two small children are pushing each other near the edge of the roof-top platform!'

34.19.1.3 *Derived Stative Propensity Verbs*

Rgyalrong employs a range of prefixes to derive stative verbs correlated with diverse *propensity* meanings.

²⁶ Inverting marking is called for owing to the marked scenario of having an identical entity serving both agent and patient roles. The inverse marker is preserved as a separate syllable before the reflexive prefix in Tshobdun (*o-ʃɐ*) and the Somang variety of Situ (*wu-ja*; Jin et al. 1958: 86); cognate reflexive markers Showu (*vʃɐ-*), and Japhug (*zyɣ-*) likewise contain an inverse marker fused with the reflexive prefix. For a different view on the morphological makeup of Rgyalrong reflexive marking, see Jacques (2010a: 263).

²⁷ Situ uses a simplex prefix *ɲa-*, cognate to *ɐ-* in the other dialects (Lin 1993: 256–7). An intransitivizing function may be attributed to this morpheme (Jacques 2010a: 262, 2012a: 212–13).

²⁸ Among the Rgyalrong dialects, reciprocal formation involves total reduplication in Situ but partial reduplication in Tshobdun, Showu, and Japhug.

The prefix *sɐ-* transforms a transitive action verb into a stative one meaning “have a propensity for V-ing (people).” This is a *subject-oriented* derivation in that the propensity is predicated of the subject argument of the source verb. As with the homophonous [+HUMAN] antipassive in *sɐ-*, the unexpressed object of this derived intransitive verb must be [+HUMAN]:

- (18) *tʃʰu?* ‘gore’
sɐ-tʃʰu? ‘have a propensity for goring (people)’
ɓʰdu ‘beat, punch’
sɐ-ɓʰdu ‘have a disposition towards beating (people)’
sə-rɲi ‘lend’
sɐ-sə-rɲi? ‘have a disposition towards lending things to others; be generous’

The prefix *wɐ-* attaches itself to both transitive and intransitive stems to derive another subject-oriented stative verb with the meaning “tend to V; be easy to V.” Examples are:

- (19) *jmət* ‘forget’
wɐ-jmət ‘tend to forget; be forgetful’
rgɐzʔ ‘age’
wɐ-rgɐzʔ ‘tend to age at an accelerated rate’
tu ‘be ripe’
wɐ-tu ‘be easy to ripen’

Two other derivational prefixes with a similar stativizing effect are attested. Both resultant intransitive verbs attribute the predicated propensity to the object argument of the source verb.

The object-oriented prefix *nwə-* applies to transitive action verbs, yielding stative verbs with the meaning “be easy to be V-ed”:

- (20) *ɓʰdu* ‘beat, punch’
nwə-ɓʰdu ‘be easy to be beaten; be weak’
rɲi ‘borrow’
nwə-rɲi ‘be easy to borrow (things) from’
mtiʔ ‘see’
nwə-mtiʔ ‘be easy to be seen (as of e.g. mushrooms)’
jmət ‘forget’
nwə-jmət ‘be easy to be forgotten’

The other object-oriented prefix *sɐ-*, possibly embodying an extended usage of the causative *sɐ-*, turns bivalent verbs with experiencer subjects into stative verbs meaning “easy (for one) to V her/him/it.” Examples are:

- (21) *mti?* 'see'
sɛ-mti 'be visible (lit. making one tend to see it)'
tsi 'understand'
sɛ-tsi 'be easy to understand (lit. making one tend to understand her/him/it)'
jmət 'forget'
sɛ-jmət 'be easily forgotten (making one tend to forget it)'

34.19.2 Category-changing Derivation

A morphology-rich language, Rgyalrong puts a heavy constraint on the distribution of word-classes in functional slots. Verbal stems typically function only as predicates, and must be nominalized to serve as clausal arguments or modifiers. Likewise, nominal stems cannot head verb phrases without first undergoing conversion to verbs by means of verbalization morphology.

Certain concepts which languages typically encode as verbs (e.g. 'be angry,' 'laugh') are conveyed by nominal roots in Rgyalrong, which must be verbalized or enter into a verbal construction to serve as clausal predicates:

- (22) *tə-múru* 'anger (nominal prefix + nominal root)'
rɛ-múru 'be angry (verbalizing prefix + nominal root)'
tə-múru ʔgu? 'get angry (N + V)'
tɐ-jmi? 'dream (nominal prefix + nominal root)'
wɛ-jmi 'dream of (verbalizing prefix *wɛ-* + nominal root)'

A small number of verbal concepts are conveyed by roots labile in word-classes, in that they directly combine with either nominal or verbal affixes with no overt marking of category transformation:

- (23) *kɐ-rət* 'to sweep (infinitive prefix + root)'
tə-rət 'sweeping (nominal prefix + root)'
kɐ-ɐfɕʰɐ 'to cough (infinitive prefix + root)'
tə-ɐfɕʰɐ 'coughing (nominal prefix + root)'

Zero-derivation is only scantily attested. More common are cases where the nominal root assumes a phonologically adjusted form of the verb root, usually Stem 2:

- (24) *kɐ-ʔge* 'to wear (infinitive prefix + root)'
tə-ʔge 'clothes (nominal prefix + accented root)'
kɐ-mtser 'to be hungry (infinitive prefix + verb root)'
tə-mtserʔ 'hunger (nominal prefix + verb Stem 2)'
kɐ-qjɔ? 'to vomit (infinitive prefix + verb root)'

<i>tə-qjɔ</i>	‘vomit (nominal prefix + verb Stem 2)’
<i>kɐ-vɐr</i>	‘to fear (infinitive prefix + verb root)’
<i>tə-verʔ</i>	‘fear (nominal prefix + verb Stem 2)’
<i>kɐ-ɐmqɛʔ</i>	‘to quarrel (infinitive prefix + verb root)’
<i>tə-ɐmqɐ</i>	‘quarrelling (nominal prefix + verb Stem 2)’

34.19.2.1 Nominalizing Derivations

Most nominalizations in Rgyalrong have a *clausal* scope and a verbal internal structure.²⁹ Action/state nominalization in *kə-/kɐ-*, for instance, produces nominalized clauses (cf. *the enemy(s) destroying the city*) rather than noun phrases (cf. *the enemy’s destruction of the city*). Nominalizing structures, however, do result in true lexical deverbal nouns (see also Jacques 2008: §2.4).

34.19.2.1.1 Bare Verb Roots Used as Nouns

A restricted set of common verbs in root form may occur directly as verbal nouns carrying a possessor prefix representing the notional subject, as seen in:

- (25) a. *jɐpe* *jə-ⁿdze* (*jə-t^hi*) *nɐ-mts^hôt*
 last.year 1PL:POSS-eat (1PL:POSS-drink) IMPV:PST-be.much
 ‘We had a lot to eat (and drink) last year.’
- b. *koʔ* *ⁿbri* *ó-nkɐ* *tʃ^hə* *ʃsət*
 this horse 3SG:POSS-walk what be.like
 ‘How is this horse’s gait?’

Another interesting usage is reported in the Japhug dialect (Jacques 2008: 106), where certain complement-taking verbs allow a complement clause headed by such possessed verbal roots:

- (26) *rɲul* *nɐ-mbi* *mu-pu-rɲó-t-a*
 money 2SG:POSS-give NEG-PFV-experience-PST-1SG
 ‘You have never given me money (lit. I have never experienced you giving me money).’

34.19.2.1.2 Deverbal Nouns in *tə-*

The nominal prefix *tə-* can attach to many verbal roots to form deverbal nouns of various functions. The verb stem selected in this formation varies depending on the dialect. In Tshobdun, Stem 2 is usually chosen, but exceptions abound. They may be lexicalized into patient nouns in some instances, such as *tə-zɐ* ‘food (nominal prefix + root

²⁹ See Sun (2012) for a full presentation of the various types of clausal nominalization in Tshobdun Rgyalrong and their applications in complementation constructions.

‘eat’) and *tə-wɛ* ‘clothes (nominal prefix + root ‘wear’)’ in the Situ dialect (Lin 1993: 267). Deverbal nouns are also used as action/state nouns. Certain phrasal idioms employ deverbal nouns of this type, such as:

- (27) *ləmu* *tə-vɛvo* *ɐsɐ-jɔ=cə*
 Lamu NMLZ-weep CONT:HTR-lift=MED
 ‘Lamu is crying to the heavens.’

In another widely attested usage, deverbal nouns in *tə-* occur in a complex verb structure headed by light verbs meaning “do,” “put,” or “throw,” illustrated by the following examples in Japhug (28a; Jacques 2008: 99) and Tshobdun (28b):

- (28) a. *tɛrte* *u-tu-tɕuβ* *pu-lát-a*
 hat 3SG:POSS-NMLZ-sew PFV-throw-1SG
 ‘I sewed the hats.’
 b. *ɐji?* *ɛ-ta* *tə-márkət* *mə-tə-tə-tɛ?*
 1SG 1SG:POSS-top NMLZ-steal NEG-IMP-2-place₃
 ‘Don’t charge me with the theft!’

34.19.2.1.3 Vestigial Nominalizing Suffix

Analysis of the Tshobdun form *tə-márkət* ‘theft’ in (28b) reveals a suffix *-t*, which reflects an ancient Proto-Sino-Tibetan nominalizing suffix **-s* still surviving as a non-productive nominalizer in the Situ and Japhug dialects (Jacques 2003), as shown by:

- (29) Situ nominalizing suffix *-s*
- | | |
|----------------------------------|--|
| <i>Verb root</i> | <i>Deverbal noun</i> |
| <i>ro</i> ‘visit (a patient)’ | <i>sa-ro-s</i> ‘gift brought by visitor’ |
| <i>tsut</i> ‘hide (information)’ | <i>ta-tsu-s</i> ‘secret’ |
| <i>rit</i> ‘sweep’ | <i>tə-ri-s ka-pa</i> ‘to do sweeping’ |

34.19.2.1.4 Participant Nominals

Participant deverbal nouns refer to entities denoted by the source verbs. Affixation of the subject (*kə-*) and object (*kɐ-*) nominalizing prefixes to Stem 1 regularly creates derived nouns denoting respectively the subject and object arguments of the verb. These include terms of occupations (e.g. *kə-qɛrɲi?* ‘hunter,’ *kə-loy* ‘herder’) and physical characteristics (e.g. *kə-nʰgi?* ‘patient,’ *kə-sɲo* ‘lunatic’), and miscellaneous common nouns such as *kə-cʰi?* ‘sugar, sweets (lit. sweet one)’ and *kɐ-ndze* ‘food (lit. that which is eaten)’ in Tshobdun, and *kə-jām* ‘sun (lit. sunny one)’ in Situ. Nouns derived

by the oblique nominalizer *sɐ-*, on the other hand, are open to instrumental (e.g. *sɐ-cuʔ* ‘key’), locational (*sɐ-loʔ* ‘pasture’), as well as temporal (*sɐ-sceʔ* ‘birthday’) readings. Another productive deverbal formation is manner nominalization in *tə-*, exemplified as follows:

- (30) *qʰoʔ* *tɛkʰu* *kɛ-ji=nəʔ*
 SEQ opium GP-plant=TOP
ɲəmə *o-tə-ji* *te-kɛ-pe* *ŋoʔ*
 maize 3SG:POSS-NMLZ:MANNER-plant IMPV-GP-do be
 ‘As for planting opium, one does it the same way as one plants maize.’

34.19.2.2 Verbalization

Denominal verbs are formed by adding a set of semantically differentiated verbalizing prefixes to noun stems.

34.19.2.2.1 Verbalizer *nə-*

This prefix, apparently the default verbalizer in Rgyalrong, applies freely to various noun stems³⁰ to create denominative verbs with a wide range of meanings:

(31)	Noun stem	Denominal verb stem
	- <i>pu</i> ‘son’	<i>nɐ-pu</i> ‘regard as one’s son’
	- <i>ŋɐ</i> ‘sun’	<i>nɛ-ŋɐ</i> ‘sunbathe’
	- <i>qɛchwet</i> ‘fox’	<i>nə-qɛchwet</i> ‘hunt for foxes’
	- <i>ciʔ</i> ‘water’	<i>nə-ciʔ</i> ‘(as of an animal) drink water’
	- <i>mɲa</i> ‘eye’	<i>nə-mɲa</i> ‘aim at’

34.19.2.2.2 Intransitive Factitive Verbalizer *rə-*

Denominal factitive verbs can be derived from a dedicated verbalizing prefix *rə-*, which turns a noun stem (N) into an intransitive verb of the meaning “make N”:

(32)	Noun stem	Denominal verb stems
	<i>ləju</i> ‘song’	<i>rə-ləju</i> ‘sing songs’
	- <i>mɛ</i> ‘labor’	<i>rɐ-mɛ</i> ‘do labor’
	- <i>pu</i> ‘child’	<i>rɐ-pu</i> ‘bear children/young animals’
	<i>ⁿdzem</i>	<i>rɐ-ⁿdzem</i> ‘build bridges; (as of a river) freeze over’

³⁰ In a related category-preserving usage, the prefix *nə-* attaches to stative verb roots (e.g. *mpʰjor* ‘be beautiful’) to derive transitive verbs meaning ‘regard as V[STATIVE]’ (e.g. *nɐ-mpʰjor* ‘regard as beautiful’).

34.19.2.2.3 *Transitive Factitive Verbalizer wə-*

This (non-alternating) verbalizing prefix derives a *transitive* verb with a factitive meaning from a nominal stem:

(33)	Noun stem	Denominal verb stem
	<i>jmiʔ</i> 'dream'	<i>wé-jmi</i> 'dream of'
	<i>riʔ</i> 'thread'	<i>wə-riʔ</i> 'thread (a needle)'
	<i>mot</i> 'gush of air'	<i>wé-mot</i> 'blow (a gush of air)'
	<i>zdeʔ</i> 'companion'	<i>wé-zde</i> 'accompany'

34.19.2.2.4 *Verbalizer ewə-*

This prefix complex, also associated with a constant semantic value, turns a noun stem (N) into an intransitive stative verb of the meaning 'be rich in N; be full of N':

(34)	Noun stem	Denominal verb stem
	<i>-lo</i> 'milk'	<i>ewə-lo</i> '(as of a cow) be productive in milk'
	<i>fet</i> 'strength'	<i>ewə-fet</i> 'be strong'
	<i>-zgre</i> 'sound'	<i>ewə-zgre</i> 'be loud'
	<i>-múru</i> 'anger'	<i>ewə-múru</i> 'be quick-tempered'
	<i>-k^hu</i> 'smoke'	<i>ewə-k^hu</i> 'be smoky (as of firewood)'

34.20 DERIVATIONAL MORPHOLOGY OF IDEOPHONES

34.20.1 Ideophone Root and Patterned Variation

Ideophones are special vocabulary items that depict vivid sensory images about sounds, shapes, colors, properties, and actions. The category-specific morphosyntax of ideophones in Rgyalrong justify their recognition as a distinct word-class (Sun 2004a; Jacques 2008: 305–14). A true ideophone in this language is analyzable into a monosyllabic *ideophone root* couched in a specific morphological pattern. The root expresses the core semantic content, whereas the formal pattern provides a modulated aspectual meaning. The following sentence illustrates a derived ideophone *pəpətsrəytsrəy*, which conveys an explicit image of not only a distinct type of sound (cracking noises made by wood, denoted by the root *tsrəy*) but also the manner of the action creating the sound (continuous, disorderly action involving multiple actants, denoted by the morphological form):

- (35) *val̥u=kə* *k^he* *pəpətsrəytsrəy* *ne-sətsəʔ*
 earthquake=ERG house IDEO PFV:TR-cause.to.do₂
 ‘The earthquake caused the house to make all kinds of cracking and squeaking noises.’

Remarkably, ideophone roots of similar meanings are typically not cognate across the dialects, and the derivational patterns are also often disparate. Nine such patterns have been noted in the Tshobdun dialect, shown in Table 34.1.

These patterns are illustrated as in (36) (see also example (35)):

Table 34.1 Ideophone alternation patterns

Pattern	Formal feature	Aktionsart value
I	<i>R</i> ideophone root by itself	semelfactive action ³¹
II	<i>RRʔ</i> reduplicated ideophone root followed by glottalization	state
III	<i>nə-RRʔ</i> reduplicated ideophone root prefixed with <i>nə-</i> and followed by glottalization	quick and continuous action
IV	<i>R-ə-R</i> reduplicated ideophone root with <i>-ə-</i> inserted in between	continuous and rhythmic action
V	<i>R-nə-R</i> reduplicated ideophone root with <i>-nə-</i> inserted in between	intermittent and rhythmic action
VI	<i>R-XəXə</i> ideophone root without coda followed by two identical syllables with coda of ideophone (X-) for onset and <i>-ə</i> for rhyme	deliberate and continuous action viewed in close-up
VII	<i>pəpə-RR</i> reduplicated ideophone root prefixed with <i>pəpə-</i>	disorderly state or action involving one or multiple participants
VIII	<i>Yopə-Yolevʔ</i> Onset of ideophone root (Y-) attached to <i>-opə</i> and <i>-ole(v)ʔ</i>	
IX	<i>Yov-Ye</i> Onset of ideophone root (Y-) attached to <i>-ov</i> and <i>-e</i>	

³¹ The prefix *p^hə-* can be added to Pattern I for emphasis.

- (36) a. Pattern I
 káxsɛ sekóʔ-q^hu jər tə-ɔʃəsətsəʔ
 leopard bushes-back IDEO PFV-make.oneself.do₂
 ‘In one motion, the leopard stealthily hid himself behind the bushes.’
- b. Pattern II
 káxsɛ sekóʔ-p^hi jərjərʔ t^hɐ-smo=cə
 leopard bushes-underside IDEO CONT:LTR-stay=MED
 ‘The leopard stayed stealthily (and unmoving) under the bushes.’
- c. Pattern III
 káxsɛ nəjərjərʔ t^he-re
 leopard IDEO PFV:downriver-go₂
 ‘The leopard ran stealthily and quickly downriver.’
- d. Pattern IV
 káxsɛ jərɐjər t^he-re
 leopard IDEO PFV:downriver-go₂
 ‘The leopard trotted rhythmically and stealthily downriver.’
- e. Pattern V
 káxsɛ jərnəjər t^he-re
 leopard IDEO PFV:downriver-go₂
 ‘The leopard moved stealthily downriver, one step after another.’
- f. Pattern VI
 káxsɛ jərərə t^he-re
 leopard IDEO PFV:downriver-go₂
 ‘The leopard moved downriver in deliberate, stealthy steps.’
- g. Patterns VIII & IX
 censənʔ nəʔosɛfsət t^hopət^holeʔ/t^hovt^he=kə
 Censen like.that IDEO=ADV
 k^hɛnaŋ le-nəre=cə
 home PFV:upriver-go.back=MED
 ‘And Censen went back home in that disorderly drunken manner.’

The inherent meaning of a given ideophone root may clash with the aspectual values of certain derivational patterns, leading to systematic gaps. For instance, the ideophone **pəpəjərjər* is non-existent, since the root *jər* describing silent stealthy movements is incompatible with Pattern VII *pəpə-RR*, which suggests noisiness and lack of order.

34.20.2 Derived Ideophone Verbs

Ideophones in Rgyalrong fill many grammatical roles. They are typically predicate-modifiers, but may also serve as resultative complements and even verbal

and nominal classifiers. To function as clausal predicates, however, they must be converted into *ideophone verbs* by adding one of three verbalizing prefixes *nə-*, *wɛ-*, and *sɛ-* to reduplicated ideophone roots. The derived ideophone verbs are fixed in form and meaning, as shown in (37) with the root *ʃkel* depicting the manner of walking with a limp:

- (37) *nə-ʃkelʃkelʔ* ‘walk about quickly and continuously with a limp’
 wɛ-ʃkélʃkel ‘limp about quickly and continuously in a small area’
 sɛ-ʃkélʃkel ‘cause to limp about quickly and continuously in a small area’

In order to produce predicative ideophones capable of covering the entire range of Aktionsart meanings found in non-verbal ideophones, one must resort to a syntactic strategy; namely, putting ideophones in a construction with ideophone-introducing verbs *kə-sətsə* and *kə-pe*, as shown in (35).³²

34.21 CONCLUSION

Derivation, coupled with compounding and incorporation, produces word-formation processes Rgyalrong draws on to enrich its lexical inventory. Perhaps more than any other language in the Sino-Tibetan family, Rgyalrong displays abundant derivational processes that shape stem forms, transform word-classes, and create derived words within a given word-class.

Noteworthy features of Rgyalrong derivational morphology include preference of prefixation over suffixation and, more strikingly, the presence of non-concatenating derivational devices, such as tonal and accentual modulations and ablaut. Likewise uncommon traits, especially from a Sino-Tibetan perspective, are the diverse derivational strategies for altering lexical categories or manipulating argument structure, and the generation of derived ideophones from phonologically well-defined roots via rule-governed Aktionsart patterns in a theme-variation relationship.

As in other languages, lexical derivation applies with varying degrees of productivity, but the more regular processes are accessible even to recent loan words, as shown by the highly visible verbalizing prefix *nə-*:

³² Details are in Sun (2004a: §2.5.2). The ideophone introducing verbs *sətsə* and *pe* occur respectively with auditory and non-auditory ideophone roots. This is a consequence of their grammaticalization from the verbs *ʃsə* ‘say’ and *pe* ‘do’, reflecting a widely observed cross-linguistic tendency (Childs 1994: 187–8; Schaefer and Egbokhare 2002).

- (38) a. *tʃaxpe* ‘robber (from Tibetan *jag.pa*)’
 nə-tʃaxpe ‘rob’
 b. *fúni* ‘sideline job (from Chinese *fùyè* 副業)’
 nə-fúni ‘do a sideline job’
 c. *kóntso* ‘government job (from Chinese *gōngzuò* 工作)’
 nə-kóntso ‘take a government job’

These, and other topics highlighted in this chapter, testify to the amazing robustness and richness of derivational morphology in Rgyalrong.

CHAPTER 35

PAMA-NYUNGAN

JANE SIMPSON

35.1 INTRODUCTION

AUSTRALIA provides an important study case for how people coin terms for new concepts, since the 1788 invasion resulted in many speech communities developing hundreds of new terms very rapidly. Hunter-gatherer peoples suddenly came into contact with people who were sedentary, who came with new animals, new weapons and technologies, who cooked and ate food in different ways, and who had a diversity of occupations and social categories to maintain their settlements. In the north of Australia, Aborigines had come into contact with sea-farers from Sulawesi and other parts of Indonesia well before 1788. New terms created by borrowing from Makassar, Bugis, and other languages have been investigated (Walker and Zorc 1981, Evans 1992), but investigating terms for new concepts formed by other means is rendered difficult by the long periods between the start of the contact with the sea-farers, the arrival of English-speaking settlers, and the linguistic study of northern languages.

The 1788 invasion immediately affected speakers of Pama-Nyungan languages in the Sydney area on the east coast of Australia. Pama-Nyungan languages¹ make up the largest language family in Australia, both in terms of the number of languages and the

¹ Languages and their sources discussed in the text and not otherwise referenced include: Arabana (Hercus 1994); Arrernte (Henderson and Dobson 1994), Mparntwe Arrernte (Wilkins 1989); Adnyamathanha (Schebeck 1998); Barnjarla (Schürmann 1844); Bundjalung (Sharpe 1995); Diyari (Austin 1981); Garrwa (Mushin 2012); Gurindji (Meakins et al. 2013); Kauria (Teichmann and Schürmann 1840); Kokata (Platt 1972, Coleman 2012); Ngardi (Cataldi n.d.); Ngiyampaa (Donaldson 1980); Nyangumarta (O'Grady 1960, Sharp 2004, Wangka Maya Pilbara Aboriginal Language Centre 2010); Paakantyi (Hercus 1993); Walmajarri (Richards and Hudson 1990); Wangkajunga (Jones 2011); Warlpiri: *Warlpiri Dictionary* refers to the version, formatted by Christopher Manning in 2009. Citation: (Laughren et al. 2007); Warumungu (Simpson 2008b); Western Desert (Raa and Woenne Green 1974); Wik-Mungkan (Kilham et al. 1986); Yarlalde (Meyer 1843); Yir Yoront (Alpher 1973, 1991); Yolngu Matha (Zorc 1986).

area covered (most of Australia, except for Tasmania and languages in the north and northwest). They share many properties, and their status as a genetic group is reasonably well established (Bown and Koch 2004). Major sub-groups of Pama-Nyungan include the well documented Arandic sub-group located in Central Australia (including Arrernte), the Western Desert languages occupying much of Western Australia and western South Australia (including Wangkajunga, Pitjantjatjara, and Kokata), the Ngumpin-Yapa sub-group (including Warlpiri, Gurindji, and possibly Warumungu), the Thura-Yura languages of southern South Australia (including Kurna, Barngarla, and Adnyamathanha), the Karnic languages (including Diyari), the Paman languages of Cape York, and various other groups including Yaralde. The number of speakers is small; the 2011 Australian Census indicates that only around 32,000 people said they spoke a language which would be classified as Pama-Nyungan, and many of those were people who identified with a language, but did not use it every day (Biddle 2012; Simpson 2008a). Only a few of the languages are still spoken by children.

General overviews of Australian languages focus mostly on Pama-Nyungan languages (Dixon 1976, 1980, 2002, Yallop 1982, 1993, Blake 1987, 1991, Koch 2007), and there is an important early survey of derivational morphology (Dixon 1976).

35.2 WORD FORMATION IN PAMA-NYUNGAN

Types of word-formation processes in Pama-Nyungan languages are affected by the syntactic word-classes of Pama-Nyungan languages, and the syntactic structures those word-classes appear in. Many Pama-Nyungan languages have two principal lexical word-classes, nominals and verbs, and some have two open classes, nominals and coverbs/preverbs, and a closed class of verb roots. Nominals are a large productive class denoting things and attributes of things or events (i.e. including some concepts that might be expressed with adverbs, verbs or adjectives in other languages). Pama-Nyungan languages vary as to whether the class of verb roots (denoting actions, thoughts, and ideas) is productive, or whether it is a small closed class with an open class of coverbs/preverbs which are loosely associated syntactically with the verb.

The fact that there are only two major word-classes in most Pama-Nyungan languages has two consequences for derivational processes. The first consequence is that the most common category-changing processes are those creating verbs and nominals. The second is that a considerable number of derivational processes change one semantic class of a word category into another semantic class of that same category; for example, changing words that denote things into words that denote properties.

Derivational processes are also affected by the prevailing syntactic structures in Pama-Nyungan languages, in particular, how grammatical relations are expressed. Grammatical relations are mostly expressed by case suffixes, and not by word order. There is a continuum between case suffixes and derivational suffixes, that is, between suffixes which express syntactic relations and suffixes which create new words. For

productive suffixes, it is often impossible to draw a line between their use in expressing syntactic relations and their use in creating new words. Second, because most Pama-Nyungan languages use word order for information structure purposes rather than for indicating grammatical functions, and most make frequent use of appositional and dislocational structures rather than having large rigidly ordered noun phrases, it is sometimes hard to tease out the preferred structures which lead to the creation of compounds, and grammaticalization leading to derivational morphology. However, most have a rich array of derivational affixes ranging from the productive and transparent (which often are hard to tell from postpositions or auxiliary verbs), to the unproductive which may be semantically opaque, or may contain frozen analogues of modern productive suffixes, or “cran” morphs.

35.3 PROCESSES

The processes of the creation of words for new concepts via derivation, compounding, and extension were discussed in a classic paper (O’Grady 1960) on Nyangumarta, a language spoken along the central west coast of Western Australia and the Great Sandy Desert. Speakers had sustained contact with English-speaking settlers from around 1864. Of O’Grady’s examples, about a quarter were borrowings from English, and most of the remainder were created by semantic extensions, derivational affixes, and compounding.²

Extension, derivational affixation, compounding, and borrowing strategies are found across Pama-Nyungan languages. Table 35.1 and the footnotes illustrate these with terms for “boot” or “shoe”³ in Pama-Nyungan languages spanning the continent: northeast (Cape York: Yir Yoront and Wik Mungkan), north (Arnhem Land: Yolngu Matha), centre and west (Ngumpin-Yapa: Walmajarri, Gurindji, and possibly Warumungu; Marrngu: Nyangumarta, Western Desert: Pitjantjatjara; Arandic: Arrernte), south (Thura-Yura: Kurna and Barngarla; Yalalde), southeast (Paakantyi).

In the Adelaide area, a number of terms for new concepts were recorded shortly after the 1836 invasion and settlement of Kurna country in an early grammar of Kurna (Teichelmann and Schürmann 1840, Amery 1993). The short period between settlement and recording, and the great distance between Adelaide and the nearest colonies (Sydney and Tasmania) mean that these terms are likely to have been locally formed and to have reflected strategies common in the local speech community. Extension is an obvious way of expressing new concepts, extending the meaning of an existing word,

² O’Grady also noted examples of modifications in the meanings of words borrowed from other indigenous languages, and onomatopoeia.

³ Some speech communities had traditional words for foot coverings, whether bark sandals to protect from hot sand (Garrwa *barlkabarlka*, feather shoes worn to hide tracks, or possibly skin shoes used for warmth, Western Desert *paltara* type of “shoe” made out of opossum skin). These are not discussed here.

Table 35.1 Words for "boot" or "shoe" in some Pama–Nyungan languages

Language/sub-group	Form	Gloss	Structure
Yir Yoront ^a	<i>thaml</i>	foot	Extension
Yolngu Matha	<i>dhapathu</i>		Borrowing from Makassan
Pitjantjatjara	<i>puuta</i>		Borrowing from English
Warumungu ^b	<i>jina-kari</i>	foot- possessive	Derivation
Nyangumarta	<i>jina-puka</i>	foot-stink	Compound
Wik–Mungkan	<i>tha' murruk</i>	<i>tha'</i> 'foot' <i>murruk</i> unglossed	Compound; second part suggested as onomatopoeic "creak" (Sutton 1995)
Paakantyi	<i>thina-puta</i>	foot- <i>boot</i>	Compound with English borrowing

^a Compare also Gurindji *jamana* 'foot, shoe' and Barngarla *bokirri* 'anything to rest the feet on.'

^b Compare also Walmajarri *jina-purru* 'foot-purposive' (in entry for *wirrpakangu*) Yolngu Matha *luku-puy* 'foot/sole-associative'; Arrernte *ingke-kenhe* 'foot-belonging.to'; Yarlalde *turn-inyeri* 'foot-belonging.to'.

or specifying it in some way, and Kaurna examples include *makki* 'ice, glass.' Many are formed by compounding: *tipogadla*⁴ 'gunpowder, matches' (*tipo* 'spark', *gadla* 'fire'), or by affixation. Affixation to nominals can be illustrated by a form **(i)yana* (possibly a form of the possessive suffix) which is added to body parts or body products to create clothing terms associated with them: *mukartiana* 'hat' (*mukarta* 'head'), *turtiana* 'jacket' (*turta* 'arm'), *tikiana* 'waistcoat' (*tiki* 'region of ribs'), *nukkeana* 'pocket-handkerchief' (*nukke* 'mucus'). New terms are also formed by adding suffixes to verbs, including reduplicated forms: *bilti-tti* 'pair of scissors' (*bilte-ndi* 'cut'), *nainmainma-tti* 'pair of pincers' (*nainma-ndi* 'to press or keep together'). These can in turn be compounded, *tipo-ngatpangatpa-tti* 'ramrod' (*tipo* 'spark', *ngatpa-ndi* 'to go down, to put into').

In Central Australia Warumungu people encountered Europeans in the late 19th century. A hundred years later (Simpson 1985), about half the words for new terms were borrowings from English or one of the indigenous creoles that had developed in the 19th century, and there were also borrowings for new terms from other indigenous languages. The most common means otherwise was to use a suffix. Like Nyangumarta, these include the possessive suffix (the most common means of creating new nominals in Warumungu), e.g. *jina-kari* 'boot' (foot-possessive), and the "having" suffix, for example *kunapa-jangu* 'Greyhound bus' (dog-having). A suffix meaning "excessive" was sometimes used: *wangangu-yilppi* 'alcoholic' (grog-excessive). Compounding was sometimes used, *kuwarta junmarn* 'rabbit' (ear long), but mainly for natural kind terms. There were also a number of loose phrasal compounds involving nominalized verbs and

⁴ The spelling used is that of Teichelmann and Schürmann (1840).

nominals with possessive suffixes, for example *miyili-kari nya.nj.alkki* (eye-possessive see nominalizer associated) ‘glasses.’

McGregor’s (2000) study of terms for the new concept “policeman” in more than a hundred Pama-Nyungan and non-Pama-Nyungan Australian languages also shows morphological processes of nominal and verbal derivation, and compounding. As in Warumungu, the most common nominal-derived terms for “policeman” use a comitative “having” suffix (West Mudburra *jikala-wurru* ‘chain-comitative’) or a genitive suffix (Kukatja *wayin-watji* ‘chain-genitive’). There were some nominal compounds (Nyangumarta *wirtirr-ngumpa* ‘stern face’). In some languages words for “policeman” were derived from verbs, and showed more variety including an agentive suffix (Jarú *wirin.gaji* ‘tie.up-agent’), a habitual suffix (Wangkajunga *karrpil-payi* ‘tie.up-habitual’), a purposive suffix (Yugambé Bundjalung *ganilih-gu* ‘tie.up-purposive’), and reduplication (Arrernte *irrkwe-nhe-irrkwe-nhe*, ‘policeman’ from *irrkwe.me* ‘take hold of’ with a nominalizer *-nhe* ‘thing associated with,’ and reduplication). There were also compounds with nominalized verbs (Martu Wangka *ngalyi-karrpilpa* ‘neck-tie habitual’). (All examples are from McGregor (2000) who gives the sources.)

We turn now to look at the strategies in more detail.

35.3.1 Compounding

O’Grady’s compounding examples in Nyangumarta include *pitapita warnku*⁵ ‘feral cat’ (forehead stone), *milimili yarnta* ‘write a letter’ (paper poke); it is unclear how closely bound they are. This uncertainty is found in other Pama-Nyungan languages. If juxtaposition of nominals is a favorite syntactic strategy, then this can lead to high usage of compounds, making it not always easy to tell whether a sequence is a word or not. As Hercus writes of Arabana:

Because of the free use of juxtaposition to mark inalienable possession there is no clear demarcation line between juxtaposed nouns and compound nouns. The simplest criterion is purely semantic: a compound is regarded as a unit. (Hercus 1994: 100)

Hercus gives examples such as *mardi-papu* hairstring-egg = ‘ball of string,’ *thidna-pirri* foot-claw = ‘toenail.’ In some languages, such as Warlpiri, there are morphological tests showing if sequences of two nominals really are a unit; the choice of ergative allomorph depends on the number of syllables in the word. The word *kuna* ‘anus, faeces’ is two syllables, and so takes the two-syllable allomorph *-ngku*: *kuna-ngku*. The word *mulyu-kuna* snout-faeces ‘Black-headed Python’ is treated as having four syllables, and so takes the polysyllabic allomorph *-rlu*: *mulyu-kuna-rlu* (Warlpiri Dictionary: *mulyu-kuna* entry).

⁵ I have changed O’Grady’s orthography into that used by the Wangka Maya Aboriginal Language Centre, and used in their Nyangumarta e-dictionary.

In Warlpiri and Warumungu, nominal compounds seem to be more common for expressing natural kind terms than artifacts. Some Pama-Nyungan languages use compounding quite extensively: Yir Yoront (Alpher 1991) appears to have a considerable number of compounds for new things, in part because of the common use of generics in syntactic structures generally: *kawn-wil* ‘alcohol’ (water sour), *kawn kapl* ‘light beer’ (water light), *may-thow* ‘bean sprouts’ (*may* generic vegetable food, *thow* shavings), *kay-mel* ‘bullet’ (*kay* implement, metal, glass, *mel* eye, point), *thaml-kay* ‘horse shoe’ (foot/shoe, metal). They can be expanded: *kalq-kay-thum-thum* ‘branding-iron’ (*kalq* spear, *kay* implement, metal, glass, *thum-thum* hot). There are compounds with deverbal nominals: *yoq+penpn pam nhin-nh* ‘chair’ (thing-flat person sit-participle: flat thing for a person to sit on), *may+waym-lhonne* ‘baking powder’ (*may* generic vegetable food, *waym* swell up, *-lhonne* complex suffix creating a participle).

Another major compounding process appears in the creation of verbs. Many Pama-Nyungan languages have a productive process of expressing action/event/state meanings by compounding a verb stem with a nominal (Koch and Hercus 2013), or with a “preverb” or “coverb.” The latter open class words occur with verbs, and may create new verb meanings compositionally or non-compositionally, as in Warlpiri (Nash 1986, Simpson 2009). These have been extensively discussed for the non-Pama-Nyungan languages Jaminjung (Schultze-Berndt 2000) and Wagiman (Wilson 1999), and see also Amberber et al. 2010). Meanings of the preverbs/coverb range widely, from adverbial meanings to goal meanings to quantificational meanings. These presumably arose from preferred structures such as [X Verb] or [N Verb] which then became sufficiently fixed in word order for idiosyncratic meanings to arise and, in time, for some of the nominals or other categories to lose their ability to appear independently of the verb. Finally verbs may be created by compounding with other verbs which add aspectual, temporal, and associated motion/path meanings (Koch 1984, Wilkins 1991, Simpson 2001) (see Section 35.4.2.4).

35.3.2 Reduplication

Reduplication in Australian languages has many functions (Fabricius 1998), and the functions depend on the morphological category and the semantic class of the element being reduplicated. The Nyangumarta word for the new concept “dress” *jina-jina* is said to be a reduplication of “foot,” deriving from a time when dresses hung down to the feet (Wangka Maya Pilbara Aboriginal Language Centre 2010). But as a process, reduplication does not appear to have been a major strategy; O’Grady only discusses it glancingly under onomatopoeia, where he notes *tingkitingki* ‘bell’ and *ngurrngurr* ‘pig.’ The latter form exemplifies the fact that reduplication is involved in forms denoting introduced animals in many languages. Table 35.2 shows reduplicated words for “pig” across a number of Pama-Nyungan languages.⁶

⁶ Other strategies for forming words denoting “pig” are compounding: Arabana *wadlhu-pakali* from *wadlhu* ‘ground’ *paka-* ‘to dig’; borrowing: Kurna and Diyari *piki*; and borrowing + classification Yir Yoront *pik*, *minh-pik*; *minh* ‘game animal.’

Table 35.2 Words for “pig” in some Pama–Nyungan languages

Language and sub-group	Form	Gloss/comment	Structure
Yolngu Matha languages	<i>bikipiki</i>	said to be from Kriol ^a	borrowing + reduplication
Yolngu Matha languages	<i>nurrng'nurrng</i> ^b		reduplication + onomatopoeia
Warumungu	<i>oink-oink</i>	recent borrowing ca. 1990 ^c	reduplication + English onomatopoeia

^a Compare *piki-piki* found across the centre of Australia in Arabana (Hercus 1994: 96), Nyangumarta, Pitjantjatjara, Walmajarri, Warlpiri, and Warumungu, as well as Arrernte *piki-piki* (Wilkins 1989: 150, 151).

^b Cited in Amery (1993: 44) as from Zorc (1986). Compare also Nyangumarta *ngurr-ngurr*; Pitjantjatjara *nguurnguurpa*; Warlpiri *nguur-nguurrpa* cf *nguur-mani* 'snore, grunt'; Arrernte *ngwerrkngge-ngwerrkngge*; Wangkajunga *ngurrngurr nguurr* 'the grunting sound of a pig' (Jones 2011: 38, 59). The distribution is similar to that of the reduplicated borrowing *piki-piki*. Another reduplicated onomatopoeic form for 'pig' is Walmajarri and Ngardi *ngangkirr-ngangkirr* cf *ngangkirr* 'grunting sound'.

^c After the death of a person the sound of whose name resembled 'piki', *oink-oink* was used instead.

The forms show considerable similarity. The fact that the unreduplicated borrowing *piki* is attested early in Kurna (Teichelmann and Schürmann 1840), and then to the north of it in Diyari, is an indication that perhaps the reduplicated forms were transmitted through speakers of creole and pidgins. This is consistent with the English word “pig” being expressed in the Katherine/Roper River Kriol dictionary as *bigibigi ~ pigi-pigi* (Lee 2004). Reduplicated forms for the names of introduced animals are also found in the English-based creoles across Australia, and *naninani* ‘goat’, *dakdak* ‘duck’ were recorded by the end of the 1850s in South Seas Jargon (Crowley 1993). Thus, when these forms appear in some Australian languages, it is uncertain as to whether they arose from reduplication by local people, or whether by borrowing of the English form or of a reduplicated creole form.

Onomatopoeia is involved in some of the reduplications: the Nyangumarta word for “pig” *ngurrngurr* represents the sounds pigs make. The onomatopoeiac reduplications are strikingly similar, but it is not yet determined whether the similarities are due to shared inheritance of onomatopoeia, or to borrowing. Traditionally, reduplication of words denoting sounds is quite common; an example from the Nyangumarta dictionary (Wangka Maya Pilbara Aboriginal Language Centre 2010) is:

- parl* noun. noise of a thump or thud. As of kangaroo giving a danger signal with its foot.
- parl-jini* verb (intransitive). he thumped.
- parl-parl*, noun. onomatopoeic for the noise of a repeated hitting. E.g. used of a Crested Pigeon’s wingtips hitting each other when taking off and flying strongly.

35.3.3 Derivation

Derivational suffixes discussed by O'Grady include:

1. A suffix *-pinti* (glossed by O'Grady as “complement of”) commonly used on nominals and nominalized verbs to create words denoting “thing associated with”: *ngarnka-pinti* ‘razor’ (*ngarnka* ‘beard’); *wirrka-na-pinti* ‘knife’ (cut-nominalizer-*pinti*).
2. A possessive suffix *-mili* on nominals and nominalized verbs (often compounded) to make words denoting things or places. None of the forms he recorded are listed in the modern Nyangumarta dictionary (Wangka Maya Pilbara Aboriginal Language Centre 2010), and one of them, *punta rapu-rapu jini-mili* ‘baking powder’ (ash swollen making-of), has now been replaced by *purnta rapu-rapu jina-pinti* using the *-pinti* suffix instead. However, the *-mili* suffix as a derivational suffix is still found in some older derived words: *wangkarru-mili* ‘spider-web’ (*wangkarru* ‘spider’).
3. A habitative suffix: *karikiti* ‘alcoholic’ (*kari* ‘bitter’ extended to alcohol) (the modern dictionary has for the same meaning a non-assimilated suffix *-kata*, *karikata*).
4. A suffix of ‘result’: *kari-ja* ‘drunk’ (alcohol-result)
5. A ‘having’ suffix: *lipa-jartiny* ‘leper’ (leprosy-having)

Many of these are found in other Pama-Nyungan languages. In the following sections, derivation is discussed first in terms of category-changing processes, and then some major semantic categories of derivation are presented.

35.4 CHANGING MORPHOLOGICAL CATEGORY

35.4.1 Creating Nominals from Verbs

In many Pama-Nyungan languages, nominals are formed from verbs via creation of participles. A widespread participle form has been reconstructed for proto Pama-Nyungan as **+ñca* (Evans 1988: 94, Alpher 2004: 107). This is illustrated in Ngiyampaa where the participle marker *-DHa:yN-* creates verbal nominals which may denote actions or actors, for example *mulari-dja:yN* ‘shivering, shiverer’ (<*mulari-y* ‘shiver’). These may be compounded with aspectual suffixes: *ñiya-l* ‘say’, *gara:-y* ‘all day’, *ñiya-gara:-dha:yN-* ‘talkative, talker.’ If the verb is transitive it may be loosely compounded with a nominal expressing the object of the verb, to create a new nominal.

- (1) *dhumba manuma-da:y mayi*
 sheep.ABS⁷ steal-CM.PART.ABS person.ABS
 ‘sheep-stealing fellow’ (Donaldson 1980: 199–200)

Forming the nominal may require adding a suffix to the nominalizing participle marker:⁸ a suffix for instrument terminates Wangkajunga *mintim-ma-nku-nja-pinti* sew-CAUS-AUG-NOMZ-INSTR ‘needle’ which is formed by borrowing a creole verb “mintim” (<mend), adding a causative transitive verb formative *-ma-*, adding the nominalizing participle marker *-nja* (along with a conjugation-specific augment homophonous with the irrealis) (Jones 2011: 169).

Depending on the language, nominalized verbs may be used for creating subordinate clauses, and, with different endings, for creating new terms, as in Warlpiri (Nash 1986, Hale et al. 1996). For example, in Diyari, the participle form of the verb can be used in combination with a tensed auxiliary or other verb, as in (2).

- (2) *kupa-kupa mindi-ṇa karṭi-ṇa waṇṭi-yi puṇa-ṇi*
 REDUP-child.ABS run-PART go.round-PART AUX-PRES hut-LOC
 ‘The children ran round the hut’ (Austin 1981: 99, ex. 113)

Suffixed with *-yiṇa* ‘habitual association’ it forms nouns denoting “animate beings who habitually carry out the action specified by the verb stem,” for example *mindi-ṇa-yiṇa* ‘runner’ (run-PART-HAB) (Austin 1981: 163–4). If the verb is transitive it can then be loosely compounded with a nominal expressing the object, for example *yawada yinṭki-ṇa-yiṇa* (word give-PART-HAB) ‘preacher.’ These loose compound structures are quite widespread; see further discussion of the constructions in Warlpiri (Simpson 2009).

35.4.2 Creating Verbs from Other Elements

In many Pama-Nyungan languages, verbs are intrinsically transitive or intransitive. Often, Pama-Nyungan languages have conjugation classes whose members are all or mostly transitive, or intransitive. Two common and related strategies for creating

⁷ 2 second person, 1 first person, ABS absolutive, AGENT, ALL allative, AUG Augment, AUX auxiliary, CAUS causative, CM conjugation marker, COMMIT comitative, DAT dative, DIM diminutive, ERG ergative, EUPH euphonious; GEN genitive, HAB habitual, IMP imperative, INCH inchoative, INSTR instrument, INTRANS intransitive, LOC locative, N noun, ND Nyangumarta dictionary, NFUT non-future, NOM nominalizer, NOMZ nominalizer, NPAST non-past, OBL oblique, PART participle; pl plural, PRES present; PRIV privative, PURP purposive, REDUP reduplication, S subject, Sg singular, TRANS transitive, YY Yir Yoront, portmanteau (no clear morpheme boundary).

⁸ This participle marker appears to be derivational rather than inflectional.

new verbs are compounding of an element to an existing verb (preverb/coverb–verb constructions), and attaching a verbalizing suffix to another element. Sometimes, zero derivation is found, for example, Jones argues for the Western Desert language Wangkajunga that a number of nominals denoting entities are transformed into verbs of the *-la* class (prototypically transitive) by simple addition of the endings of the *-la* class (which often begin with the conjugation marker *-l-*), e.g. *puyu* ‘smoke,’ *puyu-la* ‘smoke it!’ (Jones 2011: 171).

The distinction between compounding and verbalizing suffixes is not always clear. The problem is particularly acute in languages of the centre and north, where there is a tendency for verbs to be a closed class, and for the verbalizing suffixes to belong to particular conjugation classes, and sometimes to be homophonous with existing verbs. Arguments can be made for treating apparent verb derivational affixes in Warlpiri as special instances of preverb–verb constructions (Nash 1986: 42ff). Sharp makes a case in Nyangumarta for distinguishing some monosyllabic verb-creating morphemes (which she considers affixes), from other forms, which she considers to be compound constructions comparable to the Warlpiri preverb–verb constructions described by Nash.

35.4.2.1 Derivational Suffixes/Bound Verbs Creating Transitive Verbs

Across Pama-Nyungan languages there is a widespread formative *-ma-* forming verbs—mostly transitive verbs, but sometimes intransitive as well. This *-ma-* form is very widespread across Pama-Nyungan languages for “get, take,” and is reconstructed for proto Pama-Nyungan as **ma-* ‘take,’ and possibly a bound verb **-ma-* (Alpher 2004: 462). Alpher cites as cognates verbalizer uses in eastern New South Wales (NSW) (Bundjalung). Here *-ma-*, when attached to adjectives or intransitive verbs (including detransitivized verbs) means “to cause to be X/do X,” when attached to uninflected nouns means “to make X,” and when attached to nouns in the locative means “to put into/onto X” (Crowley 1978). A particularly complex example shows the *-ma-* attaching to a noun with two suffixes—where the Genitive creates a new nominal “English language” (belonging to whiteman), and the Instrumental indicates the language medium:

- (3) Wana: yirali-ŋa:-yu-ma-:.
 don’t whiteman-GEN-INSTR-CAUS-IMP
 ‘Don’t speak in English.’ (Crowley, 1978: 89 ex.26)

The *-ma-* form may vary from being the main way of forming transitive verbs from other elements (as in Warlpiri), to being one of a set of transitive verb formatives, as in Ngiyampaa and Nyangumarta.

Warlpiri has one highly productive formative which creates transitive verbs, causative *-ma-*, and a number of less productive forms. The formative differs from regular verbs in preverb–verb constructions in that the order of the element *-ma-* cannot be reversed, whereas productive preverbs can appear immediately after the verb. Added to nominals denoting attributes, the *-ma-* form creates transitive verbs denoting change of state (e.g. *yiri-ma-ni* ‘sharp-causative-nonpast’ = ‘sharpen’), as well as

creation (*karnta-ma-ni* ‘woman-causative-nonpast’ = ‘make a woman (from a possum)’). It can attach in a limited way to inflected nominals: *walya-kurra-ma-ni* (ground-allative-causative-nonpast), and to nominalized verbs: *wangka-nja-ku-wangka-nja-ku-ma-ni* (talk-nominalizer-purposive-reduplicated-causative-nonpast = ‘make speak.’ The causative *-ma-* has the same form as the verb *ma-ni* ‘to get.’ It also occurs on Kriol and English loan words, where the initial “m” of the suffix blends with the Kriol transitive suffix - “*im*”: *riiti(m)-ma-ni* ‘read,’ *payi(m)-ma-ni* ‘buy.’

The *-ma-* form is seen in Nyangumarta where it is part of a set of three main suffixes creating transitive verbs, causative *-ma-*, affective *-ji-*, and *-pi-* to which Sharp does not assign a meaning (Sharp 2004: 192ff). The range of meanings for the suffixes is to some extent lexicalized. They all derive verbs of the -RN- class (Nyangumarta has two major verb conjugation classes, -NY- which are prototypically intransitive, and -RN- which are prototypically transitive, and two minor conjugation classes). Of the three transitivity suffixes, the form *-ji-* is homophonous with a verb *ji-RN* of the same class, but Sharp (2004: 203) argues that the two forms have different syntactic properties. Affective *-ji-* attaches to English or Kriol verbs to create new verbs.

In Wangkajunga this *-ma-* form is also found, along with zero derivation (creating *-la* class verbs, the largest transitive class) and a second causative derivational suffix (Jones 2011: 168ff). The two derivational suffixes are homophonous with two monosyllabic verbs *ma-* ‘get’ and *ju-* ‘put,’ and, like them, form verbs of the *-rra* conjugation class. Causative *-ma-* attaches to nominals denoting properties, causative *-ju* often creates meanings involved with putting. Of the two, causative *-ma-* attaches most commonly to English or Kriol verbs to create new verbs generally: *wirrim-ma-* ‘wear,’ *mintim-ma-* ‘sew.’ The causative *-ju-* can be attached to English or Kriol verbs with the meaning of putting: *wirrim-junu* ‘put clothes on someone or something.’

35.4.2.2 Derivational Suffixes/Bound Verbs Creating Intransitive Verbs

In Warlpiri, the most general and productive way of forming intransitive verbs is to use the inchoative formative *-jarri-mi*: *wati-jarri-mi* man-become-nonpast: ‘become an initiated man.’ The form *-jarri-mi* can attach to an inflected noun: *walya-kurra-jarri-mi* ground-allative-become-nonpast = ‘land on the ground.’ While it does not occur as an independent verb, it can be separated from the nominal by clitics; in (4) *mata-jarri* ‘become tired’ is split up by a temporal clitic and the auxiliary *ka*.

- (4) *Wurnturu ka ya-ni, mata=lku ka jarri.*
 far PRES go-NPAST tired=now PRES become.NPAST
 ‘He goes a long way, and then gets tired.’ (Warlpiri Dictionary *-jarrimi*)

Likewise in Nyangumarta, the most general way to create intransitive verbs is to use the cognate inchoative *jarri-NY*, and, less productively, the stative *karri-NY* (homophonous with a verb meaning “want, like”). *jarri-NY* follows nominals to create changes of physiological state (*munyarri jarrinyi* mouse inchoative-nonfuture ‘became a mouse,’ *wararr jarrinyi* stand inchoative-nonfuture ‘(got) into a standing position’ (Sharp 2004: 215)).

jarri-NY is very common on nominals denoting properties. Like its Warlpiri correspondent, it can follow inflected nominals:

- (5) *ngapa-majirri jarri-nyi*
 water-PRIV INCH-NFUT
 ‘The water dried up.’ (Sharp 2004: 221)

Sharp argues that *jarri-NY* and *karri-NY* differ from the transitive suffixes, not only in being disyllabic, but also in being phonological words, rather than affixes. However, she acknowledges that, unlike free verbs, they cannot occur sentence initially. Furthermore, *jarri-NY* can also form idiomatic expressions, for example *ngarlu-ngu jarri-nyi* stomach-locative inchoative-nonfuture ‘become pregnant’ (Sharp 2004: 221), and can appear with words like *nyimurl-* that do not appear elsewhere as independent words: *nyimurl jarri-nyi* ‘dived’.⁹

In Wangkajunga, the most general way to create intransitive verbs is to use the inchoative *arri-* ~ *-rri*: *jarlu-rri* big-INCH = ‘grow,’ *puntu-rri* man-INCH ‘become a man, be initiated’ (Jones 2011: 172ff). It can follow deverbal nominals:

- (6) *Wangka-parni-rri-wa!*
 talk-PRIV-INCH-IMP
 ‘Stop talking!’ (Jones 2011: 172)

Thus in these three languages, three cognate forms are used to create intransitive verbs. In each case the element to which the *-(ja)-rri-* form attaches can be an inflected nominal. The languages differ as to whether the *-(ja)-rri-* form is clearly a suffix (as in Wangkajunga), or whether it is analyzed as a verb (albeit with restricted distribution) as in Nyangumarta, or a bound verb as in Warlpiri.

In all three languages adding the inchoative verbalizer is the favored way of creating new intransitive verbs from English loanwords, for example Warlpiri *warrki-jarri-* ‘to work’ (intransitive). These may be English verbs, but may also include English nominals, for example Nyangumarta *jarrirti jarrinyi* Saturday INCH-NFUT ‘became Saturday’ (Sharp 2004: 220). Interestingly, some speech communities seem to favor particular formatives or light verbs to create new words from English or Kriol loanwords. Of Yir Yoront, Alpher writes:

The verb **ṭa-* (R, intr) ‘do’^a can be postposed to certain forms that are not verbs in Yir Yoront to form intransitive verbs. For example, from **ṭuma* “fire” is produced *ṭumṭa* “become warm.” . . This construction is apparently of limited productivity when the first constituent is a native Yir Yoront form.

It is used productively, however, to form a Yir Yoront verb from any English one: *tanṭa* “dance,” *ṭatṭa* “start,” *meriṭa-* “marry,” *preyṭa-* “pray.” English transitive

⁹ *nyimurl-* is probably cognate with the Warumungu preverb *nyimit-* which appears in the preverb-verb combination *nyimit-wanppan/nyirrinta-* ‘nyimit-fall/put = ‘dive into/put right into water or ice’.

verbs occur in this construction with their transitivizing suffix -əm (“him” or “them,” as in Melanesian Pidgin’ cf. Hall 1943: 20–1): *prêntəm̩t̩a*- “brand,” *yúsəm̩t̩a*- “use.” The derived Yir Yoront verb is, however, intransitive, and its goal is an adverbial:

<i>mip̩</i>	<i>t̩ʷorp̩</i> ,	<i>wàp̩</i>	<i>prêntəm̩t̩â</i> .
animal	calf	who	brand

‘Who branded the calf?’

Native speakers of Yir Yoront manifest irritation when an outsider picks up a construction of this type and questions them about it. They regard it as not real Yir Yoront and not real English either, a lapse of good taste that it is politest not to notice.

^a (p. 407) When this verb is used as a free form, its goal is an adverbial: *woq* “work” in *pôlpəlt̩ t̩əŋûwəl*, *wóq̩*. “let’s do work tomorrow.” Another meaning of **t̩a*- (R, intr) is “move upwards.” (Alpher 1973: 381–2)

35.4.2.3 *Emission of Sound, Light, and Fluids*

While in many Pama-Nyungan languages affixes creating verb stems are obligatorily transitive or intransitive, some exceptions need to be mentioned. There may be systematic polysemies of some affixes as to whether they create transitive or intransitive verbs. The clearest examples come from the ways in which the concepts of emission of sound, light, and fluids are expressed. In some of the central and northern languages, the same morph is used for creating transitive verbs with a range of meanings as in the Warlpiri example in (7), and for intransitive verbs from preverbs or nominals denoting sound, light, and fluids ((8), (9)). The intransitivity is demonstrated by their having Absolutive subjects ((8), (9)), as opposed to Ergative subjects as in (7).

- (7) *Yapa-ngku ka warlu jarra-ma-ni.*
 person-ERG PRES fire tongue.of.flame-CAUS-NPAST
 ‘A person lights a fire.’ (Warlpiri Dictionary: entry for *rduul-yi-nyi*, hn481)
- (8) *ngulaju ka maarr-maarr-ma-ni wanjilypiri=ji.*
 then PRES flash-flash-EMIT-NPAST star.ABS=EUPH
 ‘then the stars twinkle.’ (Warlpiri Dictionary: entry for *maarr(pa)* “flashing”)
- (9) *Turaki=rlangu ka murturr-ma-ni.*
 car.ABS=for.example PRES bang-EMIT-NPAST
 ‘A car, for example, goes bang.’ (Warlpiri Dictionary: entry for *murturr(pa)* “bang”)

The polysemy is also found in the cognate set produced by Alpher in support of his reconstruction of proto-Pama-Nyungan **ma*- ‘take,’ which includes transitive verbalizers, lexical verbs meaning “take,” and in Mudburra, a close relative of Warlpiri, *ma-rru* ‘talk, tell’ (Alpher 2004: 462). This polysemy can be compared with the SAY/DO

polysemy of classifier verbs in non-Pama-Nyungan languages, some of which are neighbors of Mudburra (McGregor 2002: 139–61).

35.4.2.4 *Associated Motion/Path*

Verbs often appear in sequence in Pama-Nyungan languages, in semi-serial constructions in which the meaning of one of the verbs is bleached to become an aspectual, temporal, or associated motion/path meaning. These may begin as loose syntactic compounding, in which both elements are inflected for tense, then move to a situation whereby the first one is inflected with an element that can be interpreted as a participle, then move to compounding, and finally grammaticalization as a derivational suffix. The status as a derivational suffix can be seen in some instances by the case of the subject: in the Warlpiri example in (10) the verb is created by adding a nominalizing suffix *-rninja-* to the transitive verb *kati-* ‘to tread’ (which as a finite verb takes an Ergative subject), and then joining this with the inflected verb *ya-ni* ‘go-nonpast’ which elsewhere takes an Absolutive subject. The resultant verb takes an Ergative subject, suggesting that the *ya-ni* has become grammaticalized as a derivational suffix (Simpson 2001).

- (10) Karnta-ngku ka-lu-rla kati-rninja-ya-ni jintirrjintirrpa-piya-rlu.
 woman-ERG PRES-3PLS-3DAT tread-NOM-GO-NPAST willy.wagtail-like-ERG
 ‘The women move up in a step which is like a Willy Wagtail’s movements.’
 (*Warlpiri Dictionary* entry for *yurrurnturu-ka-nyi*)

25.5 CATEGORY-RETAINING AFFIXES

A number of affixes attach to nominals to form nominals. They have two striking features. First, they include semantic categories that are shared across many Pama-Nyungan languages but with different forms. Second, they often also appear with functions more like that of case markers and, like case markers, may engage in agreement. This double nature is reflected in the name “derivational case” sometimes given to them (Nash 1986, Simpson 1991).

35.5.1 Having and Lacking

Widespread semantic categories for deriving new words include morphemes denoting “having something or someone”¹⁰ and the “absence of something.” They are found across

¹⁰ The importance of the “having” morpheme was recognized in a number of papers in Dixon (1976). These papers focused mostly on its syntactic and semantic uses, denoting temporary ownership, or accompaniment (“comitative”).

Table 35.3 Terms formed with "having" suffixes in some Pama-Nyungan languages

Nyangumarta	<i>mirlka-jartiny</i> <i>jiri-jartiny</i>	'intelligent' 'echidna'	head-comitative; sharp.point-comitative
Kaurna	<i>yangarra-tidli</i>	'married'	wife-having
Warlpiri	<i>jiri-parnta</i>	'echidna'	prickle-having
Warumungu	<i>nyili-jangu</i>		
Warumungu	<i>kunapa-jangu</i>	'Greyhound bus'	dog-having [from the picture on the side of the bus]
Yir Yoront	<i>pin.ngon lon</i>	'bullock'	horn having

Pama-Nyungan languages, often glossed as Propriative/Comitative and Privative. The "having" affix in particular is widely used for creating new words, both those denoting properties, and by extension things that have those properties. Examples are given in Table 35.3.

The morphemes are often expressed as suffixes, whether tightly bound and undergo vowel harmony, as in the Warlpiri *-kurlu~-kirli* 'having', or the Warumungu *-k(k)Vpurtu* 'lacking' (which also shows an initial fortis/lenis consonant alternation associated with tight binding), or loosely bound, as in the Warumungu *-jangu* 'having' (which lacks the initial consonant alternation). In some languages these "having" and "lacking" morphemes share properties with other case suffixes, for example in taking part in agreement, as in Walmarjarri, in which the demonstrative and the noun both are marked with the "having" ending.

- (11) *njanarti-tjarti* *kuyi-tjarti* *marna* *yani*
that-having.ABS meat-having.ABS I went
'I went with that meat' (Hudson 1976: 205)¹¹

Sometimes they are classified as postpositions; Alpher treats as postpositions the Yir Yoront *lon*, a morpheme denoting 'having,' and similarly its antonym *komrr* 'without, lacking.'

The examples given above are semantically and morphologically transparent. But frozen forms that probably meant "having" at some earlier stage can be posited for some words; for example, Yir-Yoront (*minh*)-*kalq+perr* 'echidna' (Alpher 1991) contains a

¹¹ Hudson's diacritic for retroflexion has been changed to digraphs beginning with "r," and her nominative has been changed to absolutive.

classifier *minh* ‘game animal,’ a morpheme *kalq* ‘spear,’ and a synchronically unanalysable morph +*perr* which suggests the model of the words for ‘echidna’ given in Table 35.3.

While words formed with “having” are often positive, they can develop negative implications. For example, Warlpiri has a commonly used “having” morpheme *-parnta* as in *pinkirrpa-parnta* ‘winged thing’ (*pinkirrpa* ‘wing’), *kulu-parnta* ‘angry’ (*kulu* ‘anger’). But the form resulting from suffixing *-parnta* to a nominal denoting a part of the body indicates something wrong with that body part: *milpa-parnta* ‘poor sight’ (*milpa* ‘eye’).

“Lacking,” the opposite of “having,” is used mostly for temporary lacks, but can form words where the absence of a property is characteristic, for example Warlpiri *yumurru-wangu* (fur-lacking) used as a generic for reptiles and other furless creatures. Other interesting uses include understatement such as the Warlpiri *wita-wangu* (small-lacking) used for ‘very big.’ Often the resulting form has a negative implication; the Warumungu expression *kuwarta-kapurtu* ‘ear-lacking’ is used to mean ‘not paying attention, not listening, disobedient.’ This concept was recorded for the Sydney language shortly after contact in 1788: *Ngalgear, guribúni*: Ngalgiya, ear-PRIV ‘Ngalgear, you have no ears’ (Steele 2005: 261).

35.5.2 Inhabitant of, Characteristic of, Denizen of

Another semantic category often expressed by a derivational suffix is “inhabitant of, characteristic of, denizen of,” for example Diyari *maða* ‘stone, hill,’ *maða-la* ‘hill dweller’ (Austin 1981: 40). Its importance is shown by the fact that in some speech communities it is the major way of classifying natural kinds; for example Warumungu participants in language courses in the 1990s classified birds, animals, and plants by habitat as *mangkkuru warinyi* (black.soil.plain dweller), *murru warinyi* (plains dweller), *wankkarr warinyi* (floodout dweller), *purnungkurr warinyi* (swamp dweller), *julyppu warinyi* (sandhill dweller), *karlamparl warinyi* (creekside dweller).

35.5.3 Kinship

Words denoting kin relations are often morphologically complex in Australian languages, reflecting the fact that such terms are complex, involving at least a pair of people. They may relate to “having” suffixes; a common Warlpiri term for “my wife” is *ngaju-parnta* ‘me-having’ ‘(the one) who has me’ (inferred relation of spouse). Or they may have pronouns affixed to the name of the relation indicating one member of the pair. The relation may be loose, as in Ngarrindjeri, in which the possessor pronoun follows the case suffix: *napelan* (*nap-i-l=an* spouse-Sg-SgERG=1SgGEN) ‘my wife’ (Cerin 1994: 46 ex 307). Or it may be tighter, where the possessor pronoun is absorbed as part of the stem; Koch argues that the form *-ye* found on some Arandic kin-terms such as *ameye* ‘mother’ and *akngeye* ‘father’ is an old first person possessive suffix (Koch 2003). Wilkins

shows that in Mparntwe Arrernte *meye* ‘mother’ has *me-* as the stem in the Ergative form of “her/his mother” with a new possessive pronoun *-ikwe*, but the Ergative may appear before or after the pronoun, or even in both places, with no apparent meaning difference.

- (12) *m-ikwe-le* /*me-l-ikwe* /*me-l-ikwe-le*
 mother-3kinPOSS-ERG /mother-ERG-3kinPOSS /mother-ERG-3kinPOSS-ERG
 ‘her/his mother’ (Wilkins 1989: ex. 43)

As Wilkins observes, this suggests that the pronominal possessor forms such as *-ikwe* are moving from being clitics to being suffixes.

Many languages have a kinship dyadic suffix that prototypically appears on the name of a kin relation (Evans 2003a, b), for example Warumungu *karnanti-kurangi* mother-together ‘mother(s) and child(ren),’ *kapurlu-kurangi* older.sister-together ‘older sister(s) and younger sibling(s).’ While the relation is dyadic, the number of participants may be more than two, and in Central Australian Aboriginal English the morpheme is often glossed as ‘(to)gether,’ *mother-gether* ‘mother and child.’

35.5.4 Evaluative Morphemes

A Pama-Nyungan language for which there is a rich description of evaluative morphology is Ngiyampaa (Donaldson 1980). Ngiyampaa has a relatively large number of suffixes on nominals which involve the speaker’s evaluation of something, or assessment of quantity and attitude towards that quantity. Donaldson divides derivational morphology on nominals into two kinds, those that create a nominal with roughly the same denotation as the stem nominal, and those that create nominals with clearly different denotations. Among derivational morphemes that create nominals with similar denotations as the stem nominal are diminutive and augmentative suffixes. Donaldson writes:

The “littleness” of the diminutive is associated with emotional attachment and the “bigness” of the augmentative with emotional distaste. *mayinj-djul=dhi:* (person-DIM+ABS-my) is a common greeting between friends meeting after a long separation, roughly equivalent to “My dear old friend!” *wudha mugu-giran* (ear blocked-AUG+ABS) was the angry comment which closed a description of someone’s failure to treat her mother-in-law as she should: “Great deaf thing!” ... The contrast between the affective and the literal meaning can also be consciously exploited. One informant has a one pint enamel mug which she took everywhere for some months. Though it was normally referred to as *bandigan-bidi* “big pannikin,” the owner often called it *bandigan-dhul*, particularly when she had mislaid it... -DHul conveyed attachment tinged with irony. (Donaldson 1980: 102)

Derivational morphemes that express the speaker’s evaluation and create nominals with denotations different from the stem nominal include suffixes which indicate

properties. Ngiyampaa has several of these: *winydja-buwan* ‘having mud,’ *winydja-bil* ‘having a lot of mud,’ *mil mul-gir* eye pus-NASTY.WITH ‘very mattery eye,’ *muda-bidja:l* buttocks-WITH.BIG ‘big-bottomed’ (Donaldson 1980: 112–13).

Many languages have nominal forms indicating excessive involvement with something, e.g.:

Diyari	<i>yapa-kaŋti</i>	fear-EXCESSIVE	fearful, timorous one
Diyari	<i>kupurla-yiṯa</i>	wine-HABITUAL	drunkard
Warumungu	<i>yingal(i-y)ilppi</i>	fear-EXCESSIVE	coward
Warumungu	<i>wangang(u-y)ilppi</i>	alcohol-EXCESSIVE	alcoholic

Ngiyampaa has a rich array of suffixes attaching to verb stems and preceding the tense inflection. Most of these are aspectual, but one, *-guwa-y*, “indicates that the event referred to by the verb to which it is attached is emotionally affecting” (Donaldson 1980 #2992: 194). A wide range of emotions are illustrated, from complaint to pleasure:

- (13) *waruya-l-guwa-nha=nu:*
 slander-CM-PITY-PRES=2OBL
 ‘IT’S A SHAME, (they) are running you down’ (6-117)

The salience of these categories is shown by the fact that some even survive into the mixed languages and creoles that have developed since contact with English speakers. For example, speakers of Baryulgil Square Talk (a creole based on Bundjalung) use forms like “tea-*gali*” (tea-liking) for someone who drinks a lot of tea (Fraser-Knowles and Oakes 1996: 140), compare Bundjalung *ḡura:m* ‘asleep,’ *ḡura:m-gali* ‘sleepyhead’ (Crowley 1978: 36). In the West Coast of SA *gubby* ‘water, alcohol’ and *gubby-dja* ‘drunk’ have been recorded by a Kokata woman from her mother apparently as everyday talk (Coleman 2012); *-dja* was recorded earlier for Kokata as a productive “having” affix (Platt 1972). However, it is not clear that it is still productive for these West Coast speakers, but rather has survived on a form glossed as ‘drunk’ with a meaning that is highly evaluative.

ACKNOWLEDGMENTS

I thank the late Ken Hale and Barry Alpher, Luise Hercus, Harold Koch, Mary Laughren, Patrick McConvell, and David Nash among others, for much discussion of these ideas over many years.

CHAPTER 36

ATHABASKAN

KEREN RICE

ATHABASKAN languages are known for their complex morphology, especially in the domain of the verb word. These languages present challenges to typical views of word formation, perhaps the best known being that derivation and inflection do not occur in the expected order of derivation closer to the root than inflection. Other interesting issues arise as well. In this chapter, I highlight three of these, defining derivation, the nature of the lexical entry, and issues of productivity and idiosyncrasy.

I begin with an introduction to the family. I next focus on category-determining morphology, and then examine non-category changing morphology, describing the verb and some of the very interesting challenges that it raises.

36.1 AN INTRODUCTION TO ATHABASKAN LANGUAGES

Languages of the Athabaskan (sometimes referred to as “Dene” [“person”]) family are geographically widespread in North America, being spoken in areas from Alaska to the southwestern United States. The languages are usually divided into three groups based on geography, Northern Athabaskan in parts of Canada and Alaska, Apachean or Southern Athabaskan in the American southwest, and Pacific Coast Athabaskan in the states of California and Oregon. The languages are listed in footnote 1, organized by geographical area. Sources disagree on classification and on names; see Crippen¹ for discussion.

¹ Crippen, James. The Athabaskan languages <<http://www.drangle.com/~james/athabaskan/athabaskan-family.html>>.

A list of languages follows. Alternative names are given in some cases.

Some languages are extinct (e.g. Kwalhioqua-Tlatskanai); others have small numbers of speakers. Lewis et al. (2013) list Tolowa (Pacific Coast) with a population of one and Hupa (Pacific Coast) with eight speakers, Hän with a population of twelve and Koyukon with three hundred speakers. In all cases, *Ethnologue* reports decreasing numbers of speakers. Sekani is reported to have eighty speakers and Ahtna is said to have a speaker population of thirty-five. For Hare, the 2006 Canadian census identifies 190 of a population of 515 as having knowledge of the language.² At the other extreme, Golla (2007) reports that Navajo had around 120,000 speakers in 1990.

The Athabaskan languages are closely related and, for the purposes of the morphology addressed in this chapter, they are similar, although each is worthy of individual study. I draw all materials from the excellent grammars and dictionaries that are available.

36.2 CATEGORY-DETERMINING MORPHOLOGY

I title this section “category-determining morphology.” I use this term in two ways. The first is as a term for morphology that determines category from a lexical entry that is not specified for category and for morphology that changes category (e.g. denominal morphology). The second use is for the first of these on its own.

Category-determining morphology is generally considered to be straightforwardly derivational. In this section, I begin with a study of the root–stem distinction, asking whether category-determining morphology exists in the narrow sense. I then examine other category-assigning morphology.

36.2.1 Roots and Stem

In his classic work on Proto-Athabaskan verb stem variation, Leer (1979: 3) defines a root as “an underlying form, from which through regular phonological modification of the vowel nucleus and/or suffixation, the stem of a given stem set or group of stem sets may be derived.” He defines a stem as “root±modification±suffix.”

Northern: Ahtna, Dena’ina, Deg Hit’an/Deg Xinag, Koyukon, Holikachuk, Koyukon, Upper Kuskokwim, Lower Tanana, Upper Tanana, Tanacross, Gwich’in, Hän, Tsetsaut, Northern Tutchone, Southern Tutchone, Tagish, Tahltan, Kaska, Sekani/Tseke’ehne, Dunneza/Beaver, Babine–Witsuwit’en, Dakelh/Carrier, Chilcotin/Tsilhqot’in, Dëne Sųłiné/Chipewyan, Slavey, Mountain, Bearlake, Hare, Tłı̨chų Yat’ii/Dogrib, Sarcee/Tsųut’ina, Nicola

Pacific Coast: Coquille, Tututni, Galice-Applegate, Chetco-Tolowa, Hupa, Mattole, Wailaki, Kato, Kwalhioqua-Tlatskanai

Southern (Apachean): Navajo, Western Apache, Mescalero, Chiricahua, Jicarilla, Lipan, Plains

² <

Three major types of stems are generally distinguished in the Athabaskan literature, verb stems (the topic of Leer 1979), noun stems, and postposition stems. In the following discussion, I set aside postpositions.

One bit of evidence for a root/stem distinction is that a root may be shared by nouns and verbs, while the stems may differ, as in (1).³ In the column headed “noun,” the first form is non-possessed; the second form is possessed. A hyphen precedes the possessed form. In the column headed “verb,” the symbol “O” indicates that the verb is transitive (takes an object), and the symbol “G” that it takes a gender/noun class prefix; see section 36.4.6.2. The “Ø” is one of the so-called classifiers; see Sections 36.4.2 and 36.4.3.

(1) Ahtna (Kari 1990)

	<i>Root</i>	<i>Noun</i>		<i>Verb</i>	
	tsiic	tsiis, -tsiige'	'ochre, mineral used as paint, pigment, yellowish object' (7)	O+G+Ø+tsiic	'make mark on O, write O' (7)
	dzaek'	dzaex, -dzaegge'	'gum, resin, hardened pitch on outside of spruce' (169)	O+G+Ø+dzaek'	'caulk, smear O with gum, resin, pitch, glue' (169)

The noun stem differs from the related root and verb stem by a phonological process, spirantization, affecting the final consonant. The verb examples contain other material that is part of the basic lexical entry of the verb; see Section 36.4.

Some Navajo noun and verb stem sets are given in (2).

(2) Navajo (Young 2000)

<i>Noun</i>		<i>Verb</i>	
siil	'vapor, steam'	-ziil	'become warm' (perfective)
cha	'weeping'	-cha	'cry' (imperfective)
shéé'	'saliva'	-zhéé	'spit, expectorated' (perfective)

The initial fricative of the stem is voiceless in the noun and voiced in the verb by a predictable phonological alternation.

Nouns and verbs in examples like those in (1) and (2) share a root. One possibility is that the root is marked for category, with a verb stem derived from a noun or vice versa.

³ All data are transcribed as in their source.

Some entries in the Koyukon dictionary (Jetté and Jones 2000) imply this. For instance, we find *beet* ‘snare’ (noun; 2000: 88), followed by *O+G+Ø+beet* ‘loop O, weave O, sew O loosely, knit O’ (verb; 2000: 91); and *O+G+Ø+ghaa* ‘carry O on back, pack O’ (verb; 2000: 228) followed by *haat*, *-ghaale* ‘pack, backpack, burden’ (noun; 2000: 228), where it appears that the first might be considered to have a noun root and the second a verb root. Alternatively, in cases where a noun and verb share a root, it is possible that the root is category-less, with the stem having, in addition, a category. In cases where only a noun or only a verb occurs with a particular root, category appears to be inherent in the root.

36.2.2 Roots and Noun-stem-forming Suffixes

Writing on Witsuwit'en, Hargus (2007) says

In the verbal system, the distinction between root and stem is important and pervasive.... In the nominal system, the distinction between root and stem is less obvious. For some alienably possessed nouns, there is a possessed stem which can be analyzed as root+suffix, and in other cases as a lexically listed possessed stem. For most nouns,... there is no phonological difference between root and stem. (Hargus 2007: 239)

Hargus (2007) notes that synchronically, little evidence exists for a root/stem distinction in nouns in Witsuwit'en. There are, however, nominalizing suffixes reconstructed for Proto-Athabaskan with reflexes in the daughter languages. In general these suffixes are not productive: they are not freely available to create new stems in word formation and the meaning of the derived form may be idiosyncratic with respect to the base. I examine some of these suffixes in the following discussion.

A nominalizing suffix *-t* is reconstructed for Proto-Athabaskan. It is glossed ‘instrument for’ for Ahtna (Kari 1990), and in similar ways in other languages. Kari writes that this suffix attaches to roots to form instrumental nouns. He provides eleven nouns, saying that this list is fairly complete. Examples are given in (3).

(3) Ahtna (Kari 1990: 262)

tl'uut	‘rope’	tl'uu ‘bind’ + -t
xae	‘pack’	ghae ‘pack Object’ + -t
ggaa	‘speed, gait’	gh PROGRESSIVE + (y)aa ‘SG go’ + -t

Jetté and Jones (2000) give a similar set for Koyukon; they note that this is “an ancient noun formation process” that is fairly common. The suffix is *-t* or *-tl*.

(4) Koyukon (Jetté and Jones 2000: 375)

dzestl	‘fist, patella’	dzets ‘punch’+t
haat	‘pack’	ghaa ‘to pack’+t
tleet	‘pillow’	tleet ‘head’+o ‘classificatory compact object’+t

Holton (2000: 128) provides some Tanacross nouns with this suffix, stating that these nouns are monomorphemic.

One reason that nouns such as those in (3) and (4) are argued to have suffixes is that they are related to verb stems in meaning, as the examples show. Suffixation of *-t* is not, however, a productive word formation mechanism, and nouns with this suffix are listed.

Nevertheless, one can ask whether these nouns, while listed, are morphologically complex, consisting of a root plus a suffix, or whether they are morphologically simplex, as Holton suggests. In order to determine this, one might examine whether there are phonological differences between the patterning of root-final consonants and consonantal suffixes; different patterning might suggest that the historically suffixed forms remain morphologically complex. However, I know of no such evidence. Much of the decision about morphological complexity in these cases depends, I think, on assumptions about whether the noun and verb stems are semantically related and, if so, how. Thus, while noun and verb stems clearly can share a root, it is not obvious that this presents evidence for the level of category-less root synchronically.

36.2.3 More on Nouns and Category-determining Morphology

The suffixes discussed in this short section are of two types. The first is a morpheme that is called a kinship suffix. It does not determine or change category, but there is evidence in some languages that it is synchronically a suffix. The second type is a suffix that attaches to nouns to form adverbs, and is thus category changing.

Kari (1990: 66) lists an Ahtna suffix that he glosses “kinship suffix”; it appears in six words, and kinship terms exist that do not occur with it. This suffix also occurs in Koyukon: *-e* “kinship suffix” is in eleven nouns (e.g. *-daadz-e* ‘younger sister,’ *-koy-e* ‘grandchild’) (Jetté and Jones 2000: 10). Kinship terms exist without the suffix—Krauss (2000: 820) gives, for instance, *-lenh* ‘(man’s) brother-in-law (sister’s husband, wife’s brother).’ Most kinship terms take the suffix; most that do not Krauss identifies as originally not being kin terms (2000: 820). The Koyukon kinship terms also occur with another suffix that Jetté and Jones (2000: 5) give as *-aa*, *-aa*’ vocative (e.g. *-od-e* ‘(older) sister, (older female parallel) cousin’; *s-od-aa* vocative (Krauss 2000: 819); *-lenh* ‘(man’s) brother-in-law’; vocative *se-len-aa* (Krauss 2000: 820)), suggesting that kin terms with final *-e* are morphologically complex.⁴ There is no evidence that this suffix is category-determining in the narrow sense: the base is a kinship noun.

Kari (1990: 170) also lists an Ahtna suffix *-e* adverbial, noting that it creates adverbs, as in *dzaen* ‘day,’ *dzaen-e* ‘today.’ A cognate suffix occurs in Koyukon (Jetté and Jones 2000: 10)—*saanh* ‘summer,’ *saan-e* ‘in the summer,’ *dzaanh* ‘day,’ *dzaan-e* ‘in the day.’ This suffix shifts the category of a noun to an adverb. It appears to be non-productive: it is not used to create new words, occurring with a small class of items.

⁴ Nasals are devoiced (written *nh*) in final position; see Marlow (2000: lxx).

Derived nouns created by the addition of suffixes or enclitics that form relative clauses and nominalizations occur in Athabaskan languages.

Kari (1990) gives three Ahtna relativizing/nominalizing suffixes: *-(n)en* singular human relative suffix (1990: 66), *-(y)i* non-human relative suffix (1990: 67), *-ne* plural human relative suffix (1990: 67). These attach to verb phrases to form relative clauses. In addition, they form nominalizations, as in the examples in (5).⁵

- (5) Ahtna (Kari 1990)
- a. u'eɬ kotnes-en 'respected person' (literally: one who is known)
(1990: 66)
u-eɬ ko-t-n-es-en
3OBLIQUE.OBJECT-with areal-VOICE/VALENCE-know-NOMINALIZER
- cf. u'eɬ 'estnes 'I know him' (1990: 309)
u-eɬ es-t-n-es
3OBLIQUE.OBJECT-with 1SG.SUBJECT-VOICE/VALENCE-know
- b. kasighiɬ'aa-yi 'seal' (literally: that which has its head sticking up, that
which holds its head up) (1990: 68, 76)
ka-si-ghi-ɬ-'aa-yi
up-head-ASPECT-VOICE/VALENCE-linear object extends-NOMINALIZER
- cf. kadighi'aa 'it (stick) is sticking up' (1990: 75)
ka-di-ghi-'aa
up-NOUN.CLASS-ASPECT-linear.object.extends

Similar forms occur in other languages.

- (6) Koyukon (Jetté and Jones 2000: 443)
 kk'odaalneey-e 'lynx' (that which moves its hands around)
 kk'o-d-aa-neey-e
 perambulative-qualifier-ASPECT.VOICE/VALENCE-move.hand(s)-
 NOMINALIZER

⁵ See Section 36.4.3 for the terms used in the glosses.

- cf. yokko kk'odaalneeyh 'he is feeling around for it'
 yo-kko kko-d-aa-l-neeyh
 3DIRECT.OBJECT-on.the.surface.of perambulative-qualifier-
 ASPECT-VOICE/VALENCE-move.hand(s)

(7) Witsuwit'en (Hargus 2007: 267)

- bəχ-ən 'soldier'
 bəχ-ən
 go.to.war-NOMINALIZER
 cf. -bəχ 'go to war'

Holton identifies a Tanacross reflex of this suffix, -ε, attaching to a verb phrase. It is realized as voicing of the stem-final consonant or lengthening of the stem vowel; its tone is retained, creating contour tones.

(8) Tanacross (Holton 2000)

- a. na:tl̥̥:dz 'beads' lit. 'that which is blue' (2000: 139)
 cf. dɛltl̥̥:ts 'it is blue' (2000: 178)
 dɛ-l-tl̥̥:ts
 qualifier-VOICE/VALENCE-be.blue
 b. nɛldze:y 'candle, light' (2000: 135)
 ghɛldze:y 'moon' (137)
 cf. nɛldzey° 'it shines' (2000: 135)
 nɛ-l-dzey°
 qualifier-VOICE/VALENCE-shine

Holton notes that nouns are found that appear to be nominalizations by their morphology, but lack a synchronic base.

(9) Tanacross (Holton 2000: 135)

- tsukɛlahdze:y 'spider'

Some Navajo nominalizations are given in (10). These show one of the three nominalizers that Young and Morgan (1987) call enclitics, -í 'the one, the one that.'

(10) Navajo (Young and Morgan 1987: 1)

- a. chid-í 'automobile'
 cf. shichid 'it sits squatting'
 shi-chid
 ASPECT-squat (?)

- b. dilyĩh-í 'lead'
 cf. dilyĩh 'it melts, it is meltable'
 di-l-yĩh
 qualifier-VOICE/VALENCE-melt

The examples of nominalization show various degrees of semantic transparency. Phonological conditions on the suffixes are the same whether the suffixes are functioning as relative clause markers or nominalizers.

The suffixes discussed in this section are productive, being used to form relative clauses. At the same time, they occur with many nouns that, on the basis of their semantics, are lexically listed.

While nominalizations can be formed with a suffix, Hargus (2007: 269) notes that most Witsuwit'en nouns derived from verbs or verb phrases do not take a nominalizing suffix. Some examples of unsuffixed nominalizations are given in (11).

- (11) Witsuwit'en (Hargus 2007: 269–70)
 sɛłtsɛc 'pepper' literally 'it stings'
 nenli 'waterfall,' literally 'it flows (*li*) down (*ne*)'
 net'ɛχ 'spruce grouse,' literally 'it always (*ne*) lands (*t'ɛχ*)'
 nebi 'otter,' literally 'it swims (*bi*) around (*ne*)'

Nominalization is a productive word-formation process across the family, with suffixation, as in the examples in (5) through (10), as one possible way of nominalizing; alternatively nominalization need not involve a phonologically realized suffix, as in (11). In either case, the meaning may be predictable or not. Both the potential idiosyncrasy of meaning and the fact that, at least in some languages, it is not predictable whether a suffix is present or not, speak to the need to list these nominalizations.

Before leaving nominalization, it is worth pointing to a productive nominalizer in Carrier. Poser (2004) remarks that the use of nominalizing/relativizing suffixes varies between Carrier dialects, being obligatory in some, optional in others, and absent in others. A suffix of interest is -*ɔn* in its use as a locative nominalizer, as in the examples in (12).

- (12) Carrier (Poser 2004: 19)
 kɔnɔdaih-ɔn 'dance hall' (where they dance)
 ʔɔdinla-ɔn 'cupboard' (where multiple generic things are put)
 tadiz'ai-ɔn 'puddle' (where water puddles)
 hunzu-ɔn 'nice place' (where it is nice)

In terms of meaning, these range from what is probably best considered to be lexicalized (e.g. cupboard) to what is likely productive (e.g. nice place).

36.2.5 Summary: Category-determining Morphology

Category-determining morphology has not been a major focus in the Athabaskan literature. Suffixes are identified, but, beyond the relativizer/nominalizer, they do not appear to be used productively to form new lexemes. In some cases, there remains no evidence for morphological complexity of an historically complex form except if there is a transparent semantic relationship, and in this case one must ask whether semantic relationship requires a common root, or if the forms could be related in some other way.

To summarize, with respect to the noun-creating suffixes, there is a general sense that *-ʔ* is not productive, and it is questionable whether it is identified as a distinct morpheme synchronically in some languages. At the other extreme, the nominalizers are productive. At the same time, nouns with a nominalizer can be lexically listed—the semantic relationship with the base is not always transparent, and it is not predictable whether a nominalization will have an overt nominalizing suffix or not. Differences exist between languages, and corpus and lexical studies would be valuable in examining productivity and lexicalization in individual languages. In general, derivation beyond nominalization does not play a major role in the formation of new nouns; instead new nouns are usually formed through nominalization, compounding, and loan words.

36.2.6 Verb Roots and Stems

As discussed in Section 36.2.4, the major productive derivational process in noun formation is nominalization. In verbs, on the other hand, derivation is at the heart of the formation of new words. I begin an examination of the verb by investigating if category-determining morphology exists.

As discussed in Section 36.2.1, the root is generally considered to be a basic unit in Athabaskan languages, with stems formed through aspectual suffixation and/or modification. That noun and verb stems can be built from the same root suggests that, at least in these cases, the root might be unspecified for category, with the suffixes/modification adding the category, and thus being derivational, assuming that category-changing morphology is derivational.

In this section I begin by stepping back from the question of whether roots have a category, and introduce verb stem formation. This is a complex area of Athabaskan morphology. The primary reference is Leer (1979), and many of the grammars include discussion of verb stems.

In order to understand the role of stem-forming suffixes in verbs, it is necessary to introduce the topic of aspectual stem variation (e.g. Leer 1979). Both phonological and morphological factors must be taken into account in thinking about verb stem formation.

Roots in Athabaskan languages are generally divided into two categories depending on their phonological shape. Roots are usually monosyllabic, and can end with an obstruent or with a vowel or sonorant consonant.

Roots are also divided into categories on non-phonological grounds. Kari (1979), building on Golla (1970), identifies “verb theme categories,” or categories with shared semantic content and derivational potential. Kari recognizes several such categories, with names including motion, extension, successive, conversive, classificatory, and stative; see Section 36.4 and the grammars for discussion of verb theme categories in various languages.

Two types of aspectual categories are also recognized. One is imperfective, perfective, optative, and future. Broadly speaking, all verbs can occur in these categories. Each verb theme category is also associated with another set of aspectual categories; for Ahtna, Kari (1979) uses the terms momentaneous, continuative, conclusive, durative, and semelfactive, among others. These are basically Aktionsart categories.

Verb stems are generally displayed as a matrix, with the aspectual categories heading columns and the verb-theme category dependent Aktionsart categories heading rows. Consider, for instance, the Ahtna verb *ʼaa*, an open root in the motion verb theme category.

(13) Ahtna (Kari 1990: 70)

Root: 'aa 'classify compact object'	<i>Imperfective</i>	<i>Perfective</i>	<i>Future</i>	<i>Optative</i>
momentaneous	'aas	'aan	'aaʔ	'a'
continuative	'a'	'aan		
reversative	'aax	'aan	'aax	'aax

These verb stems share a root, *ʼaa*. The suffixes are transparent: *-s* occurs in the imperfective momentaneous, *-n* in the perfective, *-ʔ* in the momentaneous future, *-x* in the reversative imperfective, future, and optative, and *-'* in the continuative imperfective and momentaneous optative.

The verb stem system is complex, requiring more detailed attention than I can give it here. Hargus (2007: 517–18), in discussion of verb stem variation in Witsuwit'en, points to there being lexical idiosyncrasy, but, at the same time, “there are great similarities in stem variation across aspects,” with basic stem variation patterns for open and closed roots.

With this introduction to stem formation, I now ask whether these suffixes are derivational. (I use the term suffixation although there are not always overt suffixes.) The suffixes indicate aspectual categories of a variety of sorts, and negation in some languages, not illustrated here. Kari (1990: 38) treats the suffixes as inflection, while Hargus (2007: 461) suggests that the aspectual categories are quasi-aspectual derivational categories. Rice (2000) examines the suffixes with respect to several properties associated with inflection to determine if they are best considered to be inflectional or derivational. Properties associated with inflection include obligatoriness within their category, generality of meaning, wide range of applicability, weak possibility of fusion, and relevance to syntax (see Rice 2000: 295 for a summary). Rice (2000) concludes that the suffixes are derivational. Assuming the criteria above, several points suggest this. In terms of

generality of meaning, in Ahtna open roots, the glottal stop suffix occurs in imperfective, future, and optative continuative forms, and in imperfective and future distributive forms, while *-n* is the perfective suffix in conclusive, continuative, and distributive perfective forms, but not in all perfective forms. The persistive has a suffix *-x* in the imperfective, perfective, future, and optative; the reversative has this suffix in all but the perfective, where it has *-n*. It is thus difficult to associate form and meaning directly, since a single form is associated with an array of meanings. Within an aspectual class there is a wide range of applicability, a diagnostic of inflection, but there is not the type of generality that is usually associated with inflection.

I began this section asking whether there is category-determining morphology. The verb stem system speaks to productive derivational suffixation, at least historically, with much systematicity retained synchronically although with between-language differences. If some roots are category-less, then the aspectual suffix can be considered to determine category.

36.2.7 Summary

The topic of category-determining morphology has not been addressed in depth in the Athabaskan literature to my knowledge, beyond verb stem formation, and the discussion in this section perhaps points to some reasons why. Category-determining suffixes are proposed. With noun formation, outside of nominalization, productivity is limited, often so much that any sense of a derivational relationship may not be present and the relevant question is whether morphological complexity remains.

With verbs, aspectual stem variation exists. Although this system is reconstructable as suffixation and/or stem modification, in many languages the morphology has been obscured through phonological change. It is not clear if these suffixes add a category to a category-less root. At least in some languages, they are derivational, whether they are category-determining or not.

36.3 NON-CATEGORY DETERMINING DERIVATION

Athabaskan languages are rich in non-category determining derivation. Because this morphology does not determine category, a common major diagnostic for identifying an affix as derivational or inflectional is absent.

Issues around derivation and inflection figure prominently in the Athabaskan literature. While there is overall agreement on the status of particular affixes as inflectional or derivational, disagreements exist about whether certain affixes are derivational, inflectional, or something in-between.

Because of the complexities of understanding whether a particular element is inflectional or derivational, it is worthwhile to begin by reviewing these terms. Kibrik (1995/2005: 4) remarks that the distinction between inflection and derivation is of paramount importance for Athabaskan studies, given that “virtually everything that can be represented morphologically is so represented.” Kibrik proposes explicit definitions, following Mel’cuk (1993) in introducing a category “quasi-inflectional” in addition to inflectional and derivational.

Inflectional morphemes are those that serve to produce forms of one and the same lexeme. The set of all forms of a lexeme constitutes this lexeme’s paradigm. Derivational morphemes are those that serve to form new lexemes. Addition of a derivational morpheme gives rise to a new paradigm. (Kibrik 1995/2000: 8)

Kibrik proposes operational criteria to supplement these definitions, following linguists such as Mel’cuk (1993) and Bybee (1985).

Inflectional morphemes are obligatory (=obligatorily expressed in every word to which the given category applies) and regular (in terms of meaning, form, and distribution).

Quasi-inflectional morphemes are not obligatory and are regular.

Derivational morphemes are not obligatory and don’t have to behave regularly. (Kibrik 1995/2000: 8)

I address non-category changing morphology in nouns in Section 36.3.1, and introduce evaluative morphology in Section 36.3.2. I focus on the challenges of the verb in Section 36.4.

36.3.1 Nouns

Just as category-changing derivation in nouns is non-productive outside of nominalization, with debate about whether, in some cases, morphological complexity remains, non-category changing derivation is often unproductive, and it is not always clear whether morphological complexity is found synchronically.

Some prefixed nouns in different languages are shown in (14) through (17).

(14) Witsuwit’en (Hargus 2007: 240–1)

- a. də-
 dəbih ‘sheep’
 dəcən ‘stick, bush’
- b. wə-
 wəze ‘mountain alder’
 wədzih ‘caribou’

- (15) Tanacross (Holton: 2000: 126)
 dɛme: 'Dall sheep' (cf. -ma: 'grey')
 dɛʒa:z 'cow moose' (cf., -ʒa:z 'woman's son')
- (16) Navajo (Young and Morgan 1987: 4)
 dibé 'sheep'
 'azee' 'medicine'
- (17) Hupa (Golla 1970: 217)
 di-na' 'milkweed'

In discussion of such nouns in Witsuwit'en, Hargus (2007: 241) remarks that they are reconstructed as prefixed roots, but synchronically, some seem to be polysyllabic monomorphemes. There are patterns, but, Hargus notes, there does not appear to be synchronic evidence for morphological complexity. Many of these nouns do not seem to be morphologically compositional in meaning, and they are listed rather than productively formed. One might ask if they might consist of more than one morpheme, while being listed. Based on historical phonotactics of roots (see, for instance, Leer 1979), these forms would not qualify as single morphemes. However, phonotactic properties may have shifted over time. Just as original suffixes might no longer be identified as such, original prefixes might not be identified, with an accompanying loss in morphological complexity. While the question of how much restructuring has occurred is an interesting and important one, there might be little evidence to bring to bear, with decisions about complexity being dependent on theoretical assumptions.

36.3.2 Evaluative Morphology

Some languages have augmentative and diminutive morphemes. These items are classified variably as suffixes and words. They are used productively, often with an idiosyncratic meaning contribution, suggesting lexical word formation. However, in some languages they are ordered outside a possessive suffix that is generally thought to be inflectional. If inflection is external to derivation, this ordering makes it appear that the evaluative elements are words rather than suffixes, since as suffixes derivation would be outside of inflection. In addition, a type of agreement occurs with evaluative morphemes. Assuming that agreement is phrasal, again affix versus phrasal status comes under question. In this section I examine some of the complexities that evaluative morphology introduces.

Young and Morgan (1987: 2) identify Navajo adjectival suffixes that affix to nouns. One is an augmentative.

- (18) Navajo (Young and Morgan 1987: 2)
- | | | | |
|-------|------------------|-----------|---------|
| tsoh | 'big' | | |
| chidí | 'automobile' | chidítsoh | 'truck' |
| 'éé | 'garment, shirt' | 'éétsoh | 'coat' |

Ahtna has more than one augmentative. Kari notes that *cogh*, the cognate of the Navajo augmentative, is less frequent than *ce'e* (1990: 109); the latter he says is common (1990: 112).

(19) Ahtna (Kari 1990)

tsicogh	'sea otter' (big head) (109)
ʔuk'ece'e	'king salmon' (big salmon) (112)

Kari (1990: 112) classifies these morphemes as adjectives. He writes *ce'e* together with the stem when the meaning is idiosyncratic, as in (19), and as a separate word when the meaning is predictable (e.g. *gigi ce'e* 'big berries,' (1990: 112); *dzaen ce'e* 'the whole day' (1990: 113)).

Some Slavey augmentatives and diminutives are given in (20). Note that these have broader translations than just 'big' and 'small' and can be used to identify a cultural reference that stands out as of particular importance. For instance, the Mackenzie River is culturally the largest river, and a bald eagle is a culturally important large bird (20a).

(20) Slavey (Rice 1989: 238–41)

a. Augmentative cho

dih	'chicken'	dih-cho	'rooster'
sah	'bear'	sah-cho	'grizzly bear'
deh	'river'	deh-cho	'Mackenzie River'
whé	'star'	whé-cho	'evening star'
de'tone	'large bird'	de'tone-cho	'bald eagle'
jiye	'berry'	jiye-cho	'orange'

b. diminutive yaa/ya

sah	'bear'	sah-yaa	'bear cub'
nóda	'lynx, wild cat'	nóda-yaa	'domestic cat'
ts'oh	'fly'	ts'oh-ya	'maggot'
luge	'fish'	luge-ya	'herring'
'elá	'boat'	'elá-ya	'canoe'

Different analyses are found in the sources, with evaluative morphology treated as both suffixal and phrasal depending on the source, and even as both in the same language, as in the Ahtna examples discussed above. The lexical analysis is based on meaning. Evaluative items variably mean old/young or large/small member of a species; in addition, they identify specific characteristics of the modified noun. With animals, the combination of diminutive/augmentative with one noun can yield the size reading, with another the age reading, and with another the characteristic reading in the same language. If evaluative morphology is lexical, such patterning is expected. While Hargus (2007: 321) treats Witsuwit'en evaluative morphology as phrasal, she notes that the semantics of certain noun+adjective combinations resembles the "not completely compositional semantics" seen with certain noun compounds.

Based on idiosyncrasy of meaning then, these morphemes appear to be non-category affecting derivational suffixes. Nevertheless, there are reasons for treating evaluative morphology as phrasal.

Agreement provides one argument. Athabaskan languages have a morpheme that is generally called “areal” that occurs when a nominal subject of an intransitive, direct object of a transitive, object of a postposition, or possessor is “areal,” or refers to an area, place, situation; see, for instance, Thompson (1993).

With the Slavey diminutive, augmentative, and some other post-nominal adjectives, areal agreement occurs. The areal agreement is *go-*, emboldened in the examples in (21).

(21) Slavey (Rice 1989)

kóé	gocho	‘town, society’
house	areal-big	
dzene	godéwé	‘holiday’
day	areal-big	

The presence of areal agreement is generally interpreted as evidence that the adjective is a separate word.

Hargus (2007) identifies a Witsuwit'en category she calls post-nominal adjectival stems. The augmentative *co* falls in this class. Adjectives of this class exhibit areal agreement; the areal is *wə-*.

(22) Witsuwit'en (Hargus 2007: 321)

yəχ	wəco	‘big house’
house	areal-big	
qeyəχ	wəyez	‘small village’ (2007: 325)
village	areal-small	

Hargus treats Witsuwit'en post-nominal adjectives as phrasal rather than lexical despite the idiosyncrasies in meaning. Rice, on the other hand, takes the Slavey idiosyncrasies to suggest the lexical nature of word formation with these adjectives despite the agreement.

Morpheme order presents another argument that these constructions are phrasal. In the possessive construction, many nouns occur with a suffix that is usually called the possessive suffix. The Slavey suffix is *-é* or a high tone. With most post-nominal modifiers, the possessive is the final suffix, as in the examples in (23).

(23) Slavey (Rice 1989: 249)

a.	tu	‘water’
	tuwelé	‘soup’ (tu ‘water’ + wel ‘hot’ + <i>-é</i> POSSESSIVE)
b.	łétéh, -łétéhé	‘bread’
	łétéhgoné	‘hardtack’ (łétéh ‘bread’ + gon ‘dry’ + <i>-é</i> POSSESSIVE)

In a detailed study of the possessive suffix in Tłıchǫ Yatı̨ and Dëne Sų́ı́né, Wilhelm and Saxon (2010) argue that this suffix indicates syntactic licensing of a nominal, with its form and occurrence at least partially lexical. If the possessive is inflectional, its placement outside the adjective is expected. However, with the augmentative and the diminutive, the possessive immediately follows the stem, with the evaluative item final.

(24) Slavey (Rice 1989)

- ’ahécho ’hunting snowshoes’
 ’ah ’snowshoe’ + é POSSESSIVE + cho AUGMENTATIVE

In (24), the stem *’ah* ’snowshoe’ is followed by the possessive suffix, followed by the augmentative.

In some cases, the Slavey possessive follows the augmentative/diminutive (25a), or variation is possible (25b).

(25) Slavey (Rice 1989: 251)

	non-possessed	possessed	
a.	tłıcho	-tłıchoé	’horse’
	tłı ’dog’ + cho AUGMENTATIVE		
b.	sahcho	-sahchoé, -zahécho	’grizzly bear’
	sah ’bear’ + cho AUGMENTATIVE		

The diminutive and augmentative have interesting properties in the languages in which they have been studied. With these, and some other adjectives, the meaning of the word form is not always predictable from the meaning of the parts, suggesting that the derived words are lexically listed. At the same time, agreement and morpheme order suggest that the constructions are phrasal. Evaluative morphology thus raises questions about what it means to be a word, with some diagnostics suggesting lexical status, and others phrasal. See Lovick and Rice (forthcoming) for further discussion of evaluative morphology in the Athabaskan family.

36.4 THE VERB WORD

The verb word (henceforth the verb) is the focus of much study in Athabaskan languages. The verb is generally treated as composed of a root, followed by an aspectual suffix, and preceded by a number of prefixes, with the order of elements determined by a template.

The verb must be considered from two perspectives, word formation and morpheme ordering. In terms of word formation, three major levels are generally identified, the verb theme, or basic lexical entry; the verb base, or basic lexical entry plus derivation;

and the verb word, or base plus inflection. This statement alone does not capture the unusual nature of word formation for two reasons. First, what might be considered a single morpheme based on meaning can be split, with what are clearly morphemes between the pieces. Thus, units that are considered to have a single meaning are treated as composed of independent morphemes, with the meaning distributed over the whole rather than being compositional. Second, the ordering of morphemes does not mirror what might be predicted given the levels of word formation introduced above.

Kibrik (2012) illustrates the complexity of the Upper Kuskokwim verb.

- (26) Upper Kuskokwim (Kibrik 2012: 516–17; his glosses)

nontinghiji'eł 'bye, I will see you next time'						
no-	ni-	ti-	<u>ni-</u>	ghi-	s-	l- 'eł
again-	2SG.ACC-INC-	CL.roundish	PROG-	1SG.NOM-DETRANS.TRANS-	act:PROG	

Kibrik remarks that the word-final root ('eł), the pre-root marker (l-), and the discontinuous derivational classificatory prefix *ni-* constitute the lexical frame (verb theme, underlined in (26)), with inflectional morphemes added to that frame. Kibrik identifies the following prefixes as inflectional: the iterative (*no-* 'again, next time')⁶ and the future (marked by three prefixes *ti-* inceptive, *ghi-* progressive, root-final -ł, a feature of the root allomorph), as well as the subject and object. Kibrik notes that "the Athabaskan languages seem to have a high condensation of what is called 'rara and rarissima'" (2012: 517).

With this background, I focus on derivation and issues around it.

36.4.1 Verb Word Formation Introduced

Many prefixes in Athabaskan languages are classified as derivational. The following Ahtna examples illustrate. The verb stem and relevant derivational prefixes are bolded; these prefixes are called preverbs. Basic lexical entries (themes) are given on the first line of each set of data. The abbreviation "O" indicates that the verb is transitive and "G" that it takes a gender/noun class prefix. "+" and "#" represent boundary types.

- (27) Ahtna (Kari 1990)

a. basic lexical entry:	O+G+'aa 'handle compact object' (71)
dayini'aan	'he brought it in'
nayghi'aan	'he brought it down'
kayghi'aan	'he brought it up from below'

⁶ I do not address the iterative further in this work, but I simply note here that it is regarded as derivational by Kari (1990: 39) and as a quantifier by Rice (2000).

b.	basic lexical entry:	Ø+daetl' 'plural go' (144)
	da'snidaetl'	'we entered'
	nahghidaetl'	'they came down'
	ka'uhdaetl'	'you plural come up from below!'
c.	basic lexical entry:	O+G+Ø+t'aats' 'cut O with knife, saw, scythe' (341)
	iday'tnezt'aats'	'he cut it open'
	nayghit'aats'	'he cut it downward, he cut out the backbone'
	kadit'aas	'cut it (wood) up'
d.	basic lexical entry	O+G+!+kaatl' 'chop, split O with axe' (233)
	uday'tni!kaatl'	'chop it open!' (136)
	kaydghi!kaatl'	'he chopped it to bits'

Each verb word includes a stem, at the right edge. The verb can, in addition, include a preverb that affects the meaning, thus being the prototypical derivation. The preverb is at or near the left edge, separated from the stem by intervening material that includes clearly inflectional material such as subject and object. The preverbs illustrated include *ka-* 'up and out, up from below' with motion verbs (27a, b), *ka-* 'randomly, completely, into pieces' with the verbs of cutting and chopping (27c, d), *da-* entering, with motion verbs, *da-* opening with cutting and chopping verbs, and *na-* 'down' with motion verbs and with 'cut.'

36.4.2 Lexical Entries

Having illustrated derivational prefixes and stems, I now address the verb theme, or basic lexical entry. The examples in (27) show lexical entries consisting of a prefix called a classifier (voice/valence) and a stem; I ignore the root, and use the term stem.

A lexical entry may consist of just a stem, although such entries are often analyzed with a null voice/valence (classifier), indicated as Ø directly before the stem. Example (28) shows that several verb words can be built on the Ahtna theme 'compact object is in position' (Kari 1990: 70). Derivational prefixes are emboldened.

- (28) Ahtna (Kari 1990)
- | | | |
|-----------------------|---------|--|
| basic lexical entry: | G+Ø+'aa | 'compact object is in position' (1990: 70) |
| z- 'aan | | 'it is in position' (1990: 70) |
| ta-z- 'aan | | 'it is in the water' (1990: 70) |
| ta- | | 'water' (1990: 324) |
| da-z- 'aan | | 'it is up (on a shelf)' (1990: 70) |
| da- | | 'up, at an elevation' (1990: 137) |
| da-d-ghi- 'aan | | 'it is hanging (on a nail)' (1990: 70) |
| da#d | | 'hanging up' 1990: (137) |

Other positional verbs occur with these prefixes. In (29), 'sheathlike object is in position,' the voice/valence (classifier) prefix is *ɬ-*. This lexical entry thus has two overt elements.

- (29) Ahtna (Kari 1990: 319)
- | | | |
|---------------------|---------------|----------------------------------|
| basic lexical entry | <i>ɬ-taan</i> | 'it (e.g. knife, gun) is there' |
| da-d-ghi- | <i>ɬ-taan</i> | 'it (e.g. quiver) is hanging up' |
| da-de-z- | <i>taan</i> | 'it is up on a shelf' |

The lexical entry can include additional prefixes. (30) has the stem *-nen* and prefixes *d-* (qualifier) and *l-* (voice/valence).

- (30) Ahtna (Kari 1990)
- | | |
|---------------------|---|
| <i>d+l+nén</i> | 'compact object moves independently, falls' (1990: 299) |
| <i>ni-de-l-nén</i> | 'it (car) stopped' (1990: 299) |
| <i>ni-</i> | 'terminative, to a point' (1990: 311) |
| <i>na-de-l-nén</i> | 'it (key) fell down' (1990: 299) |
| <i>na-</i> | 'down' (1990: 291) |
| <i>sta-de-l-nén</i> | 'it got lost, dropped' (1990: 299) |
| <i>s+ta</i> | 'off, away, lost' 1990: (325) |

The example in (31) has the stem *-tsaa* and the prefix *da-*, forming the lexical entry, 'singular dies.'

- (31) Ahtna (Kari 1990: 367)
- | | |
|------------------------------|-----------------------|
| a. <i>da#tsaa</i> | 'SG dies' |
| b. <i>da-z-tsaan</i> | 'he died' |
| c. da-cin'-de-l-tsaan | 'he pretended to die' |

The example in (31c) includes a complex derivational element as well: the prefix 'pretend' is given as *cin'#d+l* (Kari 1990: 117).

Slavey examples of discontinuities are given in (32). The verb 'be scared, afraid' includes the stem *-ji* and the prefix *n-*; it has two morphemes, but is one unit semantically. In (32b), the stem is *-ndeh* (or variant), and the prefix *go-* areal (Sections 36.3.2, 36.4.6.2) is required on the meaning 'talk.'

- (32) Slavey (Rice 1989, Howard 1990)
- | | |
|-------------------|---|
| a. <i>n-ji</i> | 'be scared, afraid' |
| <i>ne-h-ji</i> | 'I am afraid, scared' |
| <i>go-ne-ji</i> | 'it (situation) is dangerous, frightening' (Howard 1990: 256) |
| <i>yé-n-aa-ji</i> | 's/he was afraid of it, feared it' (Howard 1990: 256) |

b. go-de-ndeh	‘s/he talks’ (Rice 1990: 720)
tì-gó-nì-nde	‘s/he started talking’ (Rice 1990: 720)
èshí-go-ì-nde	‘s/he almost talked’ (Rice 1990: 720)
ètsò-go-ì-nde	‘s/he got lost, off track in talking’ (Rice 1990: 720)
ká-go-yì-nde	‘s/he started talking, talked out’ (Rice 1990: 720)
èht’áh-no- go-nì-ndeh	‘s/he stammered, stuttered, got blocked in talking’ (Howard 1990: 365)
‘eté-go-ge-a-nde	‘they told each other stories, talked with one another, conversed’ (Howard 1990: 365)
go-dé-nde	‘s/he talked, spoke, started speaking, spoke up’ (Howard 1990: 366)
go-taa-go-ì-ndeh	‘s/he preached, gave a sermon to them’ (Howard 1990: 367)
zhá-go-ì-nde	‘s/he talked much, talked a lot’ (Howard 1990: 367)

The lexical entry can include more than one morpheme, with the core meaning distributed over the entry as a whole.

Why are lexical entries such as those in (32) considered to be morphologically complex rather than single morphemes? This is because of the discontinuities—in the surface form, the elements of the lexical entry need not be adjacent. For instance, in ‘I am afraid’ (32a), the prefix *n-* and the stem *-ji* are separated by the first person singular subject *h-*; in ‘they told stories’ (32b), *go-* and *-nde* are separated by *ge-* third person plural subject and *a-* perfective. While an infixation analysis might be proposed, it is difficult to identify phonological criteria that would position morphemes correctly. Thus, the lexical entry is often idiomatic, consisting of more than one morpheme in Athabaskan terminology.

We have seen two aspects of non-category changing morphology in verbs. First, there is rich derivational morphology, with numerous bases being derivable from a theme (basic lexical entry) through the addition of derivational prefixes. Second, the theme may be complex, or morphologically discontinuous, with pieces distributed throughout the verb; and with idiomatic meaning rather than compositional meaning.⁷

36.4.3 Templates, Predictability, and Idiosyncrasy

I now introduce the second aspect of word formation, the template. In the Athabaskan literature, the ordering of morphemes within the verb is generally considered to be

⁷ Another reason to consider discontinuous themes as morphologically complex concerns phonological properties. Languages often show differences in phonology depending on whether a sequence is morphologically simple or complex. This is, to my knowledge, not the case in Athabaskan languages. For example, the complex morphophonology involving voice/valence does not depend on whether this morpheme is transparent in function or lexically listed (thematic).

accomplished by a template (see, for instance, Kari 1989, Rice 2000). A typical (simplified) template is given in (33).

- (33) preverb # quantifier # incorporate # pronominal zone + qualifier zone + aspect zone + 1, 2 subject + voice/valence [classifier] + stem

The chart in (34), based on Hargus (2007: 458), shows the prefix template for Witsuwit'en, indicating what can function only productively and what is both productive and potentially part of a basic lexical entry (thematic).

(34)

	Productive	Thematic
Voice/valence	√	√
Inner subject	√	
Tense/conjugation/negative zone	√	
Qualifier zone	√	√
Pronominal zone	√	√
Distributive	√	√
Incorporated root	√	√
Inceptive	√	
Negative	√	
Multiple	√	√
Iterative	√	√
Preverb (postposition/adverbial) zone	√	√

Focusing on preverbs, Hargus notes that 16% of Witsuwit'en verbs (sample 830) have thematic preverbs. A few examples are given in (35).

- (35) Witsuwit'en (Hargus 2007: 456)

c'a#lih 'gamble'
de#ye 'sg., du. die'

We have seen derivational prefixes that combine with themes compositionally as well as prefixes and stems with no discernible independent meanings, forming a theme. The picture is more complex: while the combination of a preverb with some themes yields a compositional meaning, in some cases the meaning that it derives is idiosyncratic. For instance, Navajo has a prefix *'i-/a-*—that Young and Morgan (1987) gloss as ‘away, out of sight.’ This meaning is often clear. However, combined with the theme ‘laugh’ (35d), the meaning is not clearly compositional. One can imagine how its meaning arose, but it appears to be idiomatic. See Smith (2000) for discussion.

- (36) Navajo (Young and Morgan (YM) 1987, Young, Morgan, and Midgette (YMM) 1992, Smith 2000)

'i-/a- ‘away, out of sight’

- a. *'i-i-geeh* ‘to flow way out of sight, to subside (flood)’ (YM 1987: 473)
-geeh ‘move swiftly and freely through the air or over a surface, in a flowing manner’ (primary entry *-go*) (YMM 1992: 213)
- b. *'i-il-dóóh* ‘to sail away, to float away through the air (a cloud, balloon)’ (YM 1987: 475)
dóóh ‘wafting, drifting, floating through air’ (YMM 1992: 142)
- c. *'i-is-máás* ‘to roll it away (globular object, as a barrel, boulder)’ (YM 1987: 479)
máás ‘globular or cylindrical shape moves by rolling’ (YMM 1992: 396)
- d. *'i-ish-dlóóh* ‘I laugh, laugh myself to death, laugh and laugh’ (YM 1987: 481)
dlóóh ‘laugh’ (YMM 1992: 156)

Thus, a particular preverb can combine with a theme compositionally and the same preverb can combine with a theme in an idiosyncratic way.

The compositional/non-compositional patterning is particularly evident when one considers the elements that are traditionally called classifiers, and more recently identified with terms such as voice, valence, and transitivity, glossed here as follows. In the examples in (37), for the productive function I use the glosses “causative” and “middle.” For the non-productive function, I use “voice/valence” as a cover term. There are four such elements, $\emptyset$, *d*, *ʔ*, and *l*. These can function productively, as the Koyukon examples in (37) illustrate.

- (37) Koyukon (Jetté and Jones 2000)

$\mathfrak{z}$ - CAUSATIVE

- a. $\emptyset$ *lebaats* ‘it got stewed, boiled; he got scalded’ (2000: 84)
le-baats
 ASPECT-cook.by.boiling

- ‡ yeetlbaats 'he cooked it by boiling it, scaled it' (2000: 84)
 ye-tl-baats
 3OBJECT-CAUSATIVE-cook.by.boiling
- b. Ø gheetsaah 'he cried' (2000: 618)
 ghee-tsaah
 ASPECT-cry
- ‡ deketl'e ełtseh 'he makes his younger brother cry' (2000: 618)
 de-ketl'e eł-tseh
 REFLEXIVE-younger.brother CAUSATIVE-cry
- D- MIDDLE
- c. Ø yeghonh 'he is making them' (2000: 103)
 ye-ghonh
 3OBJECT-make.PLURAL.objects
- D deghonh 'they are made' (2000: 103)
 de-ghonh
 MIDDLE-make.plural.objects
- d. ‡ yeneef'aanh 'she's looking at it' (2000: 250)
 ye-nee-‡-aanh
 3OBJECT-qualifier-VOICE/VALENCE-look.at
- 1 hodenle'aanh 'she sees herself' (2000: 250)
 ho-de-n-le-'aanh
 REFLEXIVE-REFLEXIVE-qualifier-MIDDLE-look.at

In addition, each verb theme includes one of these elements. For instance, the Koyukon theme 'singular, dual swim' has Ø, and the theme 'rain' has ‡- (Axelrod 1998: 41), with no apparent motivation. As Axelrod (1998: 44) says, these prefixes "are assigned lexically where their relationship to the verb root appears arbitrary, and they may be altered via a secondary layer of transitivity-altering derivations." In the first case, the prefix is arbitrary; in the second, it is principled. Axelrod (1998: 55) concludes that these elements are "both obligatory and general *and* they have the kind of semantics we associate with lexical or derivational items—they are capable of significantly altering the conceptual content of the verb stem they accompany. They represent the ghost of what was once a productive grammatical and lexical system, which is now losing grammatical function, and merging its lexical, semantic character with the semantics of the stem." Kibrik (2012: 508) treats these prefixes as sometimes derivational and sometimes inflectional, with the derivational indicators more or less semantically motivated and sometimes completely idiosyncratic.

As we have seen, a hallmark of Athabaskan morphology is that what is a single item in terms of form and position often has a dual nature, sometimes combining with other elements in a compositional way and sometimes having no apparent independent meaning.

36.4.4 Derivational Potential

In this section I focus on derivational potential. Not all verb themes have the same derivational potential. Rather themes are viewed as grouping into categories called verb theme categories, with different categories having different derivational potential. At the largest level, verb themes are divided between active and neuter. For Witsuwit'en, Hargus (2007) distinguishes motion, activity, and onomatopoeic themes within the active themes, and subdivides motion and activity themes, distinguishing the following active theme subcategories.

- (38) Witsuwit'en verb theme categories, active themes (Hargus 2007: 556)

<i>category</i>	<i>theme subcategory</i>
motion	classificatory motion, (ordinary) motion
activity	successive, conversive, operative
onomatopoeic	

Verb theme categories are based on shared semantic characteristics, choice of aspects, derivational potential, and properties such as inceptive morphology. For instance, motion themes occur with a range of preverbs. Some are given in (39) for the Slavey verb 'singular, dual go' with the verb stem *-ʔa*. The preverbs are emboldened.

- (39) Slavey (Fort Good Hope dialect; Rice 1989)

'arats ʔedeʔa	'1,2 go home, start off home'
'ejits ʔedeʔa	'1,2 run'
'araejits ʔedeʔa	'1,2 run home'
káts ʔedeʔa	'1,2 climb out'
fíts ʔedeʔa	'1,2 go into fire'
líts ʔeʔa	'1,2 stop'
ríts'í ʔa	'1,2 arrive'
rát ʔeʔa	'1,2 go down'
tjts ʔedeʔa	'1,2 go out'
ts'elits ʔeʔa	'1,2 lose, go away'
tets'í ʔa	'1,2 go into water'

Classificatory motion themes (handling verbs) have similar derivational potential. The examples in (40) show a handling verb (*'a* 'handle default object') with some of the same preverbs as the going verb in (39).

- (40) Slavey (Fort Good Hope dialect; Rice 1989)

'arats ʔede'a	'bring home default object'
fíts ʔede'a	'put default object in fire'

ríts'į'a	'put, place, bury, set default object'
ráts'ě'a	'take down default object'
ts'elits'ě'a	'lose default object'
tets'į'a	'put default object in water'

Themes in the operative verb theme category (e.g. talk, laugh, cry, chew, sing, thaw, sharpen) occur with some of the derivational prefixes that occur with motion category themes.⁸

(41) Slavey (Rice 1989)

ká'eyǰlo	's/he burst out laughing' (1989: 720)
rírísé	's/he stopped crying' (1989: 720)

There are also derivational prefixes that occur only with operative themes. Overall operative themes have more limited derivational potential than motion themes.

While themes within a theme category take a characteristic set of derivational prefixes, there is variation among themes within a category; some are prototypical and others differ by taking more or fewer prefixes; see Axelrod (1993) and Hargus (2007) in particular.

36.4.5 Derivation and Aspect

The theme category system is important to understanding the aspect system of Athabaskan languages. As discussed in Section 36.2.6, Athabaskan languages are typically considered to have two major types of aspectual categories. One type, imperfective, perfective, future, and optative, is usually considered to be inflectional, and I do not discuss it. The other type, called Aktionsart in Section 36.2.6 and also called conjugation, is important in the theme category system. In particular, each theme category has a signature Aktionsart marker, and which aspects a verb theme can occur in depends on its theme category.

As illustrated for Witsuwit'en ((38) above), activity verbs are often divided into three major categories, generally called successive, conversive, and operative. Axelrod (1993) provides the chart in (42) showing aspectual derivations for these categories in Koyukon.

⁸ Hargus (2007) notes that in Witsuwit'en, there may be differences of meaning in how a derivational prefix combines with a theme. For instance, the prefix *ni#* with a motion verb translates 'stop and stay,' while with classificatory verbs it translates 'down on surface, to ground.' A careful study of the meanings of morphemes might show that the meaning difference has to do with the core meanings of the items involved, and is predictable.

(42) Koyukon aspectual derivations (Axelrod 1993: 124)

		Verb theme category		
		<i>operative</i>	<i>conversive</i>	<i>successive</i>
aspect	semelfactive			√
	consecutive			√
	conclusive		√	√
	durative	√	√	√
	momentaneous	√	√	√

Axelrod (1993: 125) writes that the Koyukon successive refers to activities occurring as single abrupt instances (semelfactive) versus a series of discrete instances (consecutive) versus duration of those instances such that they are no longer distinct (durative); conversive themes distinguish activities focusing on their end products (conclusive) or the process towards that end product (durative). Operative verb themes refer to a process (durative) or a punctual or terminative view of that process (momentaneous). While there can be differences in detail in how this system functions in the different languages, there is general agreement that items in a theme category have much in common in terms of derivational potential both in terms of potential aspects and in terms of derivational prefixes.⁹

36.4.6 Other Affixes

I have largely focused on stems and preverbs and their derivational potential. Preverbs are uniformly described as derivational. In this section I introduce some other verbal prefixes, examining some for which there is debate about their status as derivational or inflectional.

36.4.6.1 *Inceptive*

Some prefixes might seem to be inflectional, but are argued to be derivational; others might seem to be derivational, but are sometimes argued to be quasi-inflectional (fully regular, not involved in creating new lexemes; not obligatory).

The inceptive indicates the beginning of an event, and might be expected to be inflectional. Hargus (2007), while not explicit about its class, notes that the Witsuwit'en inceptive prefix *t-* occurs regularly with underived motion themes to give the meaning 'about to, start to, leaving' (Hargus 2007: 394). This suggests that it is inflectional, being regular

⁹ Many derivational prefixes play a role in the aspect system, with a particular derivational prefix selecting a particular Aktionsart marker. See Kari (1979), Axelrod (1993), Rice (2000) and the grammars and dictionaries for details.

and obligatory within its class. With classificatory motion verbs (verbs of handling), it can mean ‘start to take, take, steal.’ The meaning ‘steal,’ seems to create a new lexeme, suggesting derivational status.

Holden (2012) notes that in Dene Sų́łiné the inceptive occurs with motion verbs, with the fully predictable meaning ‘to set off V-ing’ or ‘to V-away.’ The form in (43a) includes *te-* inceptive, while (43b) is a periphrastic inceptive.

- (43) Dene Sų́łiné (Holden 2012)
- a. *tegha*
te-gha
 INCEPTIVE-SG.WALK
 ‘he takes off’
 - b. *ghegal héjá*
ghe-gal *héjá*
 PROGRESSIVE-SG.WALK INCEPTIVE
 ‘she starts walking’

Holden describes the inceptive prefix as giving a slightly phraseologized meaning—one expects the meaning ‘start V-ing,’ but it is ‘to set off V-ing’ or ‘to V away.’ The particle *héjá*, in (43b), contributes the meaning ‘start V-ing’ that is usually associated with inceptives. Holden concludes that the inceptive prefix is derivational. Thus, while the term inceptive might suggest inflection, this item is generally treated as derivation.¹⁰

36.4.6.2 Noun Class Markers

In Athabaskan languages nouns may be divided into classes based on properties of animacy, shape, and number. The noun class markers fall into a class of items called qualifiers. The areal, introduced in Section 36.3.2, is functionally also a type of noun class marker, although it is considered to be part of the pronominal zone rather than the qualifier zone based on its position and other factors; see (34).

The areal occurs if there is an areal patient subject of an intransitive or object of a transitive verb, object of a postposition, or possessor.

- (44) Witsuwitén (Hargus 2007: 261; glosses from source; *wə-* ‘areal’)
- | | |
|------------------------|--------------------|
| <i>sʔaq wəniłtʔəxʷ</i> | ‘it stung my back’ |
| <i>s-taq</i> | <i>wəniłtʔəxʷ</i> |
| my back | it stung area |

The prefix *n-* can occur when the patient subject of an intransitive verb or the object of a transitive verb is in the class of round objects.

¹⁰ The inceptive prefix also occurs as part of future marking; Kibrik (1995/2005) classifies this use as quasi-inflectional.

(45) Witsuwit'en (Hargus 2007: 263)

- a. sqaq ncay 'the child is big'
 sqaq ncay
 child be.big
- bənin nincay 'her face is big'
 bə-nin ni-ncay
 3.POSSESSOR-face NOUN.CLASS-be.big
- b. yəʔaʔ 'she's eating it'
 lemədec yənəʔaʔ 'she's eating the potato'
 lemədec yə-nə-ʔaʔ
 potato 3OBJECT-NOUN.CLASS-eat

The prefix *d-* occurs with several subclasses: oral noise, plants and plant by-products, large heavy objects, under syntactic conditions like those of *n-*.

(46) Witsuwit'en (Hargus 2007: 264–5)

- a. həgi 'it's drying'
 tsəz dəgi 'the firewood is drying out'
 tsəz də-gi
 firewood NOUN.CLASS-dry
- b. tsə ʔet səʔay 'a rock is there'
 tsə ʔet sə-ʔay
 rock there ASPECT-object.is.located
- tsə ʔet dizʔay 'a (heavy) rock is there'
 tsə ʔet di-z-ʔay
 rock there NOUN.CLASS-ASPECT-object-is-located

The prefixes *d-* and *n-* raise issues around their status as derivational or inflectional.

Nouns in the Ahtna and Koyukon dictionaries are marked for noun class.

(47) Ahtna (Kari 1990: 53)

- Ø tsic'uuts' z'aan 'a hat is there'
 ts'uuts' z-'aan
 hat ASPECT-object.is.located
- d- c'ecene' dez'aan 'a stump is there'
 c'ecene' de-z-'aan
 stump NOUN.CLASS-ASPECT-object.is.located
- n- gigi nez'aan 'a berry is there'
 gigi ne-z-'aan
 berry NOUN.CLASS-ASPECT-object.is.located

ko- areal	hwnax kuz'aan	'a house is there'
	hwnax	ku-z-'aan
	house	AREAL-ASPECT.object.is.located

Kari treats these prefixes as a type of derivation, as does Kibrik (2012: 497); Hargus calls them agreement markers in Witsuwit'en, without committing to whether they are inflection or derivation. Holden (2012: 569) suggests that in Dëne Sųłiné they show traces of a more complex inflectional system, and treats them as fossilized inflection. Holden further remarks that they are elements with a visible origin, rather than part of a synchronic agreement system.

If these prefixes represent agreement, one might think that they are inflectional. Why treat them as derivational or fossilized? There are several reasons. First, not all verbs enter into the agreement system. For instance, Kari (1990: 54) notes that the Ahtna verb 'buy Object' does not mark gender, while 'move compact object quickly' does. Holden (2012: 569) concludes that the majority of Dene Sųłiné verbs do not take agreement markers, and those that do must be learned. With inflection, one might expect all verbs, or a well-defined subset of verbs, to be part of the system, but this does not appear to be so. At the same time, it is important to note that, to my knowledge, there are no systematic studies of which verbs take these markers and which do not, so it is difficult to know if the verbs which occur with gender markers are a well-defined class or not. Kari (1990: 33–4) notes that Ahtna noun gender class was studied mainly with one speaker of each dialect, and not all nouns were examined.

Second, noun class markers are not obligatory. The most detailed discussion that I know of is Poser (2004) on Carrier. Poser identifies three prefixes, areal (*xw-*), *d-*, and *n-*. He notes that some verbs have four forms, with each noun class marker and with no marker. Some have only three—for instance most color terms do not have a *d-* form that is distinct from the form without a marker, but 'white' does. The following illustrates the idiosyncrasies. The noun *tl'uł* 'rope' occurs with *d-* with some verbs.

(48) Carrier (Poser 2004)

tl'uł dinca	'the rope is thick'
tl'uł di-nca	
rope	NOUN.CLASS-be.thick
*tl'uł nca	

However, with other verbs, including *dakw* 'short,' *d-* is not allowed with *tl'uł* as subject.

(49) Carrier (Poser 2004)

tl'uł ndakw	'the rope is short'
tl'uł ndakw	
rope	be.short
*tl'uł dindakw	

The verb *dɔkw* can occur with *d-* with other nouns.

- (50) Carrier (Poser 2004)

dzihtel dɔɔkw	‘the board is short’
dzihtel	di-ndɔkw
board	NOUN.CLASS-be.short

Poser concludes that these markers are not obligatorily associated with particular nouns.

Third, elements that are identical in form to noun class markers may be lexicalized with a verb. For instance, Hargus (2007: 385) notes that *d-* marks agreement in Witsuwit'en, occurring with nouns of plant parts and products. However, the *d-* in the verb in (51) is part of the lexical entry; it may have originated from the plant/wood function, but it always occurs with this verb.

- (51) Witsuwit'en (Hargus 2007)

wədiq'et	‘it's burning along’	d-q'an	‘fire goes’

Hargus treats *d-* in some verbs as derivational rather than as agreement.

- (52) Witsuwit'en (Hargus 2007: 385)

Ø-'a	‘handle compact O’	tseydin'ay	‘he put it in the fire’

The prefix *n-* is similar. While *n-* is an agreement prefix for round objects, it can occur as part of the lexical entry (Section 36.4.2), as in (53).

- (53) Witsuwit'en (Hargus 2007: 389)

n-i-ʔiʔ	‘sneak’
niʔet	‘he's sneaking by’

The areal is more robust as agreement, but can be a required part of the basic lexical entry.

- (54) Witsuwit'en (Hargus 2007: 392)

ne#w-D-tsət	‘shrink’

These items raise several issues. One is what constitutes a single morpheme. Researchers have, for instance, identified several *d*'s (see Section 36.4.7 for discussion). The noun class marker is sorted out from amongst them based on its distribution: it occurs when certain nouns are present in particular syntactic relations with another item. In other cases, a phonologically identical element in the same position is part of the basic lexical entry. These forms show characteristics of inflection and derivation both. Further lexical

work of the type done by Poser on Carrier would be helpful in coming to an understanding of these elements in individual languages.

36.4.6.3 *Incorporation*

Many Athabaskan languages exhibit noun incorporation. Some Upper Kuskokwim examples are given in (55).

- (55) Upper Kuskokwim (Kibrik 1995/2005)
- a. ho-~~ɬ~~ats-ghisdlo 'I took some dirt out'
 ɬats 'dirt'
 - b. ho-tsa-ghisdlo 'I took some rocks out'
 tsa 'rock'

Kibrik (1995/2005) says that incorporation is likely quasi-inflectional, writing that it is regular and non-obligatory. He notes that "the lack of an incorporate does not bear any specific meaning and thus cannot be considered the zero member of an opposition," adding that incorporates are added to an already inflected verb and thus do not create new lexemes.

Holden (2012), on the other hand, working within a similar framework to Kibrik, proposes that incorporation is an historical or etymological process rather than a synchronic derivational process, suggesting that verbs with incorporated nouns are lexically listed. Kari (1990) assumes that incorporation is derivational.

36.4.6.4 *Summary*

Why is there disagreement?¹¹ Different analysts are familiar with different languages, and may extend claims based on one language to the whole family. The definitions of derivation and inflection are not standardized, with different assumptions about their nature. Different data sets illustrating what is assumed to be the same morpheme may be considered in different languages. What is identified as a single morpheme by one may be different to what is called a single morpheme by another.

The issue of what is derivational is complex in the Athabaskan literature. While for the most part there is agreement, disagreements exist as well.

36.4.7 The Issue of Homophony

I close with another thorny issue, homophony. A particular phonological shape is often treated as having an array of meanings. For instance, Young and Morgan (1987: 37–8) provide thirteen entries for the Navajo qualifier *d-*; a subset is given in (56).

¹¹ See Section 36.4.3 for discussion of voice/valence (classifiers).

- (56) di-1 thematic: component of verb themes involving arms and legs
 di-2 thematic: component of verb themes concerned with relinquishment, relaxation, reduction or addition, freeing, opening or closing
 di-3 thematic: generic classifier for elongated shape, extension into time or space, and inceptive movement along
 di-6 thematic: relates to stomach, food, mouth, oral noise, noise in general, food, odor

Other *di*-s include one involving pain, one relating to color, and one relating to sound and hearing.

Kari (1990) provides similar entries for Ahtna. He identifies a *d*- with no meaning in many verb themes, and one in prefix derivational strings without an independent meaning. He also has a *d*- indicating gender (with seven gender classes), a *d*- marking benefactive, and a few others.

Kari (1990) recognizes two clear Ahtna preverbs of the shape *na*-, one meaning 'down' and the other iterative. In the Koyukon dictionary (Jetté and Jones 2000) not only is there *no*- downward and *no*- iterative (subentries iterative, again, another, circle), there is also *no*- 'across.'

Data such as these raise interesting questions. Given the same form, is it appropriate to identify distinct meanings, as is done in the Athabaskan literature? What criteria are appropriate to identify homophonous forms as distinct morphemes? Templatic position might be helpful: the iterative and 'down' occur in different templatic positions. Athabaskanists also rely on combinatorial possibilities to tease morphemes apart. For instance, the iterative generally occurs with the D voice/valence (classifier) with intransitives, while 'down' does not. Homophones might differ in required Aktionsart/conjugation—for instance, Koyukon 'down' occurs with *gh*- in the momentaneous perfective, while 'across' takes *n*- in the momentaneous perfective. Another homophonous prefix, *no*- reversative, also requires the qualifier *de*- (*no#de*+). Derivational potential might differ, with forms recognized as distinct prefixes combining with verbs of different verb theme categories. Finally, different prefixes might choose different stem forms. For instance, with the Koyukon theme *Ø-daatl* 'plural go,' *no*- 'down' takes the imperfective momentaneous stem *-daat*, while the imperfective reversative stem is *-deł* (Jetté and Jones 2000: 116).

The treatment of the Athabaskan verb as involving, by and large, a string of meaningful elements may be challenged by the rampant homophony. Careful studies of semantics might help to sort out some of these complexities.

36.5 SUMMARY

Athabaskan languages pose numerous challenges. Category-determining morphology shows little productivity, and questions arise as to whether historically morphologically

complex forms remain complex synchronically, or whether they have been restructured to be monomorphemic, creating major shifts in the canonical forms allowed. With non-category changing verbal morphology, one and the same form can be part of the basic lexical entry or function derivationally, creating a lexical item whose meaning might be compositional or idiosyncratic. The rampant homophony introduces questions about criteria for determining what an independent morpheme is on synchronic grounds. Certain items are variously analyzed as inflectional, derivational, or quasi-inflectional, with what is called a single morpheme sometimes analyzed as residing in more than one category in the same language.

From a diachronic perspective, Athabaskan morphology shows ongoing lexicalization, with what was once productive becoming frozen. At the same time, systematic, regular patterns remain. In the study of these languages, some linguists focus on systematicity, treating idiosyncrasy as idiomatic. Others focus on historical developments and how the verb word has come to be as it is. Reconciling the perspectives continues to be a challenge, and, with respect to derivation, the languages raise questions as to how well defined this concept is. I have attempted to identify several areas in which there is disagreement, and to raise questions for further study.

CHAPTER 37

ESKIMO-ALEUT

ALANA JOHNS

THE Eskimo-Aleut language family (also called Eskaleut) consists of three main divisions into the Unangax (Aleut), Yupik and Inuit languages. The latter two are more closely related and together form a grouping that has been called Eskimo.¹ The language family extends from Siberia to Kalaallit Nunaat (the land of the Greenlanders—Greenland). It is therefore indigenous language within four countries, the Russian Federation, the USA, Canada, and Kalallit Nunaat.² The members of this population live in the northern regions of each respective country; Kalaallit Nunaat is entirely northern. Originally more nomadic, people now inhabit communities near the coastal regions, usually above or sometimes just below the tree line.

There are estimated to be 96,891 speakers of Eskimo-Aleut (Lewis 2009). Vitality varies. In a number of regions, either the language or the local dialect is moribund, that is, the chances of it surviving are low. The Unangax language has a total of only 155 speakers, Yupik languages have 11,900, and the Inuit Language has 73,644 (Alaska Native Language Center). Statistics are approximate and it is important to remember that individual communities may vary widely in their language strength. In many areas the language is still very strong. This strength is clear in Kalaallit Nunaat, where everyone speaks one of the three dialects (Kalaallisut, Inuktun, and Tunumiisut). There are many communities in Canada where the Inuit language is dominant and vital. For example, in Arviat, Nunavut, of 1,915 people who claim Aboriginal identity (which would be Inuit), 1,910 claim knowledge of Aboriginal language(s), and 94.5% of those with Aboriginal identity speak an Aboriginal language most often at home (Statistics Canada 2006). This means the Inuit language is used in the home throughout this community. In contrast, in another Canadian community, Ulukhaktok, NWT, of 370 people who claim Aboriginal

¹ The term “Eskimo” is not a word from this language family but is a word used in English to refer to the people, etymology unknown, but often thought to be Algonquian. In 2010 the Inuit Circumpolar Council passed a resolution that the term “Inuit” be used in research. There also are more regional terms, e.g. Kalaallit “Greenlanders.”

² Kalaallit Nunaat is autonomous, but not fully, from Denmark.

identity (Inuit), 190 claim knowledge of an Aboriginal language and 10.8% of those who claim Aboriginal identity speak an Aboriginal language most often at home (Statistics Canada 2006). This means that the Inuit language is much less spoken throughout this community. In these two communities we see two very different language scenarios, both of which are observed in other communities across the Arctic.

Areas where the Eskimo-Aleut language holds the status of an official language include Kalaallit Nunaat, the territory of Nunavut in Canada, and the Northwest Territories of Canada.

Some important references for the language family as a whole include Dorais (2010) for a general overview of the Inuit language including sound systems, inflection, literacy, current status, extensive bibliography, etc., especially for Canadian dialects. Fortescue et al. (2010) is the second edition of the *Comparative Eskimo Dictionary*, an enormously useful compendium of reconstructed roots and derivational morphology with cognates in various dialects. One of the most thoroughly described dialects is Kalaallisut (West Greenlandic), and Fortescue (1984) is the best source, although unfortunately out of print. A good reference for Tunumiisut (East Greenlandic) is Tersis (2008). For Unangax,³ Bergsland (1994, 1997) is the most important body of work, and for the Yupik languages there are the excellent resources of Woodbury (1981), Jacobson (1984, 1995), and de Reuse (1994).

Although Unangax, Yupik, and the Inuit language are the three main divisions within the language family, each of these has subdivisions of its own. Yupik consists of three languages, Central Alaska Yup'ik, Siberian Yupik, and Alutiiq, while the Inuit language is usually considered to be one language with four main dialect groups (i) Alaskan Inupiaq; (ii) Western Canadian or W; (iii) Eastern Canadian or E; and (iv) Greenlandic. There are numerous dialects within these four groups and intelligibility across dialects has never been fully examined. Unangax has two main dialect groups, Eastern and Atkan (Western) (Bergsland 1997). For more detailed information about varieties within these main groups, see Dorais (2010) and Fortescue et al. (2010).

37.1 WORD FORMATION IN ESKIMO-ALEUT

The Eskimo-Aleut word has the general schema of root-(affixes)-inflection-(enclitic).⁴ Whereas the noun consists of one obligatory categorized root plus case, verbs consist of an obligatory categorized root plus a complex inflection (one of a number of grammatical moods combined with person and agreement). Affixes, however, are the elements which can expand in number in the middle of the word, creating the long words which

³ The actual term for the language is Unangam Tunuu but Unangax is used in English.

⁴ Unangax is the only language in the family with auxiliaries (see Bergsland 1997).

provide the basis for describing languages as polysynthetic.⁵ An example from Central Siberian Yupik is shown in (1), adapted from de Reuse (1994: 25), with deleted segments in parentheses.

- (1) angyagh-(gh)lla-ng(e)-yug-tuq=lu
 boat-big-acquire-want.to-indic.3s-also
 'also, he wants to acquire a big boat'

In this example the root on the left is *angyagh* 'boat' and the inflection towards the right *-tuq* consists of the indicative mood marker plus third-person singular. The enclitic *-lu* 'also' follows the inflection. There are three affixes internal to this word.

Derivational morphology in Eskimo-Aleut consists of the set of these affixes. This class of elements has variously been termed postbases, affixes, suffixes, infixes, etc. I will use the term postbase, which is widely used, although no label is entirely adequate. Postbases, as derivational morphology, (i) modify/extend the stem, (ii) change the category of the stem, or (iii) modify the valency or argument structure of the stem through reduction or addition. As Fortescue (1992: 247) notes:

Much of what is inflectional (or indeed lexical) in more analytic languages is clearly derivational in languages like Eskimo, where, for instance, suffixes for tense and epistemic modality have inflection-like semantics and appear just before the "real," paradigmatic/obligatory inflectional endings for mood, person, and number.

There is long standing and general consensus on this point, illustrated by the statement in Jacobson (1984: 423): "A Yup'ik *postbase* is a derivational suffix similar to English suffixes such as '-ment' in 'argument,'" or by Smith's (1978a) title for the list of these morphemes "A Survey of the Derivational Postbases of Labrador Inuttut (Eskimo)."

One exception is de Reuse (2000, 2006), who considers productive members of this set of elements not to be derivational but instead to be PNC (productive non-inflectional concatenation) elements.⁶ Lack of productivity should not be a defining property of derivational morphology. Since there is no principled way to delineate among this group of morphemes to isolate a sub-class that could be categorized as derivational morphology, I will consider the set of postbases to be derivational morphology. As we will see, some members are equivalent to derivational morphemes in more familiar languages, while others seem surprising, posing challenges for both our understanding of derivational morphology and syntax.

Enclitics will not be considered here, since they follow inflection.

⁵ Affixes are suffixal in nature; there is only one prefix in the Inuit language *ta-* which reorients deictics from the point of view of the speaker to that of another, e.g. *maani* 'here (speaker)' vs. *ta-maani* 'here (listener)', i.e. 'over there (by you)'.

⁶ Woodbury (1981) reserves the term derivational for category changing or valency changing postbases.

37.1.1 Numbers of Postbases and Productivity

Given that the set of derivational morphemes underlies the polysynthetic nature of this language family, it is not surprising that these elements are numerous; they are usually found listed either in a separate section within dictionaries of the language (e.g. Jacobson 1984) or in a lexicon of their own (e.g. Harper 1979). The important information about them concerns what each morpheme means, which categories it can attach to, whether there is any category change, etc. and what type of morphophonological effect occurs to the left as it attaches to the stem. Bergsland (1997: 105) states that in Unangax, there are about 570 postbases, with only 175 productive. de Reuse (1994: 76–7) compares the number of Central Alaska Yup'ik postbases in Jacobson (1984)—556—with the number in Central Siberian Yupik—547. He points out that many of the postbases in Central Siberian Yupik are not fully productive and suggests that Central Alaska Yup'ik has fewer non-productive postbases, such that the latter has significantly more productive postbases. Sadock (2003: 3) reports that there are approximately 500 derivational affixes in Kalaallisut. Estimating how many postbases there are is complicated by a number of factors, including whether non-productive postbases should be counted, determining whether certain sequences of postbases constitute an independent postbase, and even whether postbases in different positions within the word should be counted separately.⁷

Non-productive postbases are those which cannot be freely added or occur in only a few contexts. An example of this in Central Alaskan Yup'ik is *-tar-* 'to fetch, go to get, or gather N (not food) from nature' as in (2) from Jacobson (1984: 566). de Reuse (1994: 75) also describes the same Siberian Yupik postbase as non-productive.

- (2) *equg-tar-tuq*
 wood-collect-intr.3s
 'He is collecting firewood'

Here it would be useful to see examples of non-permitted nominal roots. Is it that this postbase selects only for mass-type, desirable products of nature? If so, this would logically limit its application, and it could be compared with the English verb *harvest*, which would not normally be used with stones, books, etc. Alternatively, it may indeed be an unproductive historical relic. In Yup'ik this postbase attaches to nominals meaning water, wood, grass, ice, and its cognates in other languages in the family attach to mussels, berries, snow, and blood (see for example Briggs et al. in progress).

Sequences of postbases are challenging for analysis. They may involve issues of lexicalization, restricted distribution or even frequency. Lexicalization is used to describe sequences where the meaning is not simply the sum of the parts. Fortescue (1980) describes the Kalaallisut postbase *-vvigi-* 'trans(itive)' as a lexicalized sequence which he says descends

⁷ This important issue is independent of the needs of dictionary makers, who may choose to posit two entries, even where a more abstract linguistic analysis would have one.

from *-vik* ‘place’ and *‘gi’* ‘have as.’ He says that this postbase only adds an object to a verbal base, and that no other postbase may intervene between the two.⁸ It would be interesting to find out if there is additional evidence for considering this type of element as one postbase (see discussion of cognates of this postbase in Section 37.2.1.2).

Bergsland (1997: 105) describes an Unangax postbase which is a composite, that is, lexicalized. The postbase *-yugaaġ-* ‘to V for a while; to V a little’ is itself a rare postbase. It is composed of the more rare *-yuug-* ‘a little’ plus the unknown *-aaġ-*. Woodbury (1981: 329) describes Central Yup’ik postbases which cannot directly combine with any root but must combine with another postbase to form a complex postbase, for example *-knaggar*, which can appear as an intensifier to other postbases: *-ta-knagga-irute* ‘for there to be absolutely no more N at subject/localis’ (*-ta* ‘N at subject/localis’); *-ng:irute* ‘to run out of N’).

Fortescue (1980: 263) makes an additional distinction among postbase sequences based on frequency. He labels as semilexicalized very common collocations, whose meaning is transparent, and which need not necessarily co-occur. One example is the Kalaallisut postbase *-qquuqi-* ‘undoubtedly; must have’ which consists of *-qqur-* ‘undoubtedly; must have’ followed by the intensifier *-qi-*. Frequency of co-occurrence is not accepted by others as a basis for a separate entry. de Reuse (1994: 69) says it is not clear what the term semilexicalization means even though “Every student of Eskimo postbases is struck by the semantics of certain combinations occurring again and again...” Fortescue (1980) maintains that this factor is related to psychologically “real” word forms, and warns against overanalysis. The delineation of postbase sequences is an understudied issue for this language, and my own observation is that the majority of these postbases are not readily perceptible by fluent speakers without training or extensive discussion of them.

A final issue concerns the length of words. Given that the properties of the postbases allow for recursion (see discussion in Section 37.2), words are in principle infinite in length; nevertheless most report a general maximum of six postbases within a word (c.f. Fortescue 1980, Bergsland 1997: 105).⁹ Longer words are possible and speakers themselves enjoy (re)creating them as a novelty. Here is a particularly long Kalaallisut word from Fortescue (1983: 97) with nine postbases.

- (3) aliikusirsuillammassuraanirartassagalarpaali
 aliklut-lirsur-i-llammak-ssuaq-u-nirar-tar-ssa-galar-paat-li
 entertainment-provide-SEMI.TRANS.-one.good.at-great-be-say-REPET.-will-sure.
 but-_{3P/3S}-however
 ‘However, they will say that he is a great entertainer, but ... (e.g. we know otherwise)’

⁸ Similarly Bergsland (1994) describes the Aleut postbase *-ġiiluġ-* ‘place for having or holding’ as a composite with its own entry, even though he says it might be a sequence of *-ġi-* ‘have as’ and *-aluġ-* ‘place for -ing.’

⁹ de Reuse (1994: 54) reports that Central Siberian Yupik usually has shorter words than Central Alaska Yup’ik and that there are usually two or three postbases. Jacobson (1984: 423) does not give us an average for Central Alaska Yup’ik but says that there are “sometimes as many as half a dozen, but infrequently more than that.”

37.1.2 Scope and Position of Affixes

The meaning of most postbases scopes over the material to which they attach (Kleinschmidt 1851, Smith 1978b, Fortescue 1980, Woodbury 1981),¹⁰ starting from the end of the word (right to left). We can see both these properties in the following Unangax examples from Bergsland (1997: 106), morpheme glosses added.

- (4) a. hla- $\hat{x}$ txin kuri-za-qali-ku- $\hat{x}$
 boy-ABS. REFL.3S smoke-HAB.-start
 ‘The boy has started to smoke (habitually), has become a smoker’
 b. qilam txin kuri-qali-za- $\hat{x}$
 in.morning REFL.3S smoke-start-HAB-
 ‘He usually starts smoking in the morning’

Here we see that the order of the postbases *-za-qali-* gives the meaning ‘start to habitually X’, while the reverse order *-qali-za-* give the meaning ‘usually starts to X.’

As a result the starting in (4b) is part of the recurring event, but is presumed to have happened only once in (4a). Kleinschmidt (1851: 110) notes that the ordering is the reverse of that we find in European languages. This inverse ordering is typical of what Cinque (1999: 53) describes as “non-closing” or stacking suffixes.

Crucially, Kleinschmidt (1851), Fortescue (1980), and Woodbury (1981) point out that scope alone does not provide an explanation for all postbase orderings. For example, there are some postbases which must be in a fixed order with respect to each other, that is, their ordering cannot be permuted. For example, Fortescue (1980) posits a set of postbases (his V_s), a group consisting of tense, negation, and epistemic modality, which must follow other verbal postbases. This suggests there are fixed hierarchical relations within the word.

Further evidence against a generalized scope rule is that in degree relations, the element on the right does not scope over the entire constituent to its left, but only the immediate element to the left, as in the Unangax example from Bergsland (1997: 105) [my glosses].

- (5) aqa-yukach-aasaada-nan a $\hat{x}$ ta-kun
 come-for.a.long.time-very-SUB.3PL appear-3P.
 ‘They apparently spent a very long time coming (here).’

Here *-aasaada-* ‘very’ applies only to *-yukat-* ‘for a long time.’ This suggests there are subconstituents within the word.

¹⁰ Both Smith (1978b) and Fortescue (1980) consider the question of whether generative rules are needed if scopal relations do the work. This issue is still debated in the work of Cinque (1999) and Ernst (2002).

Finally some of what Fortescue (1980) calls lexicalized postbases sometimes show an ordering that is unexpected, given the meaning of the components. Consider the two possible interpretations of the Kalaallisut example in (6) from Fortescue (1984: 43).

- (6) niri-qqu-nngil-aatigut
 eat-tell-not-3P/1P.INDIC.
 ‘They (i) forbade us to eat / (ii) didn’t tell us to eat’

Observe that the negative is to the right of the postbase meaning ‘tell.’ Right to left scope would lead us to expect that this example would mean only the second of the two possible interpretations, that is, there was no communication. Yet the first interpretation is also possible, where negation does not scope over the telling but only over the eating.

37.1.3 Types of Morphological Processes

The general pattern of morphological attachment is simple affixation, or concatenation, but there is often some morphophonological effect either on the final part of the stem or on both the final part of the stem and the beginning of the postbase. The effect of attachment of individual postbases is not predictable within any single variety,¹¹ and properties of cognates can vary across the languages and dialects. Each postbase must carry the information which determines the effect of its combination with preceding material. Two of the main patterns are that a postbase will either (a) directly attach to a stem or (b) that it will delete an immediately preceding consonant on a stem. A typical example of each is seen in these Uummarmiut (W) examples from Lowe 1984) in (7).

- (7) a. qaluk-tuq-tuq¹²
 fish-consume-INTR.PART.3S
 ‘He is eating fish’ c.f. qaluk ‘fish’
 b. tautu-tua-qamiung
 see-every.time-CAUS.3R/3S
 ‘Every time he sees him’ c.f. tautuk-tuq ‘He sees’; c.f. tuaq ‘every time’

In (7a) the *-tuq-* morpheme meaning ‘consume’ does not delete the final consonant of the preceding nominal *qaluk* ‘fish’ on the left. In (7b), the morpheme *-tuaq-* ‘every time’

¹¹ The only predictable variety is Labrador Inuttut, which has lost underlying codas in all roots and almost all postbases, with the exception of one or two postbases, e.g. *-niaC-* ‘near future’ where C indicates a consonant which will take its value through regressive assimilation with the following inflectional suffix; *-niak-Kunga* 1s indicative vs. *-niat-tuk* 3s participial, where the sequence kK is underlying /qq/ which is affricated (Dresher and Johns 1996).

¹² Although this looks like reduplication it is not; the intr.part. can be inflected for person, e.g. *-tunga* intr.part.1s but the postbase meaning ‘consume’ cannot. Moreover the postbase *-tuq-* appears before other inflections which do not resemble it.

deletes the final consonant of the verb stem *tautuk*- ‘see’; at the same time the causative mood morpheme (underlying form *-gamiung*) assimilates and coalesces with the final uvular stop of *-tuaq*- (making the morpheme boundary less clear).

Fortescue (1992) argues that these morphophonemic effects are distinct from simple agglutination, and that overall they enhance the saliency of individual morphemes, which helps maintain polysynthesis.¹³ Another important observation in Fortescue (1992) is that postbases do not show any evidence of having been grammaticalized from lexical items. A new root may result through lexicalization over an existing root plus a postbase (Fortescue 1980), and postbases lexicalize among themselves, but no new postbases are created from root material.

37.1.3.1 *Non-concatenative Processes*

One of the exceptions to the general pattern of straightforward affixation found throughout this language family are the so-called replacive postbases. These involve a type of affixation which triggers the deletion of the entire final syllable of the preceding stem, including also the initial consonant of the postbase. In the Inuit language, this also involves the gemination of the penultimate consonant of the stem. The replacive pattern is said to be a non-productive but frequent pattern in Kalallisut (Rischel 1974: 191–7). The general schema is shown in (8a), and a Kalaallisut example from Rischel (1974: 192) is shown in (8b), where we see an affix with both patterns, the second being the replacive version.

- (8) a. C₁V₁C₂ + C₃V₂...->....C₁C₁V₂....
 b. ipu-li^r-ppaa ~ ippi^r-ppa
 shaft-provide.with-TR.INDIC._{3S/3S}
 “(he) provides it with a shaft”

Where an affix allows the replacive pattern, it is said to be optional and the postbases that trigger it begin with /l/ or /n/ (Fortescue et al. 2010). In some cases it leads to doublets, and the non-composite meaning is associated with the replacive form. An Inupiaq example from Kaplan (1981: 252) is shown in (9).

- (9) a. iñu-liq-suq
 person-provide.with-DECL._{3S}
 ‘is provided with people’ [literal]
 b. iññiqsuq (with replacive -liq-)
 ‘has visitors’

¹³ Another interesting morphophonological issue concerning postbases is the fact that many postbases begin with consonant clusters. The phonotactics of the languages generally do not permit consonant cluster onsets, thus the extra consonant must be syllabified in the coda of the preceding syllable of the stem. An example from the Utkuhtsalingmiut dialect (W) from Briggs et al. (in progress) is: *-qpaluk*- ‘looks like’ *ataata-ga-qpaluk-tuq* father-poss.1s-looks.like-part.3s ‘He looks like my father.’ This is syllabified *a-taa-ta-gaq-pa-luk-tuq*.

The fact that replacive morphology is linked to the less compositional meanings appears similar to recent discussion of diminutive derivation where non-compositional diminutives are analyzed as being closer to the root (De Belder et al. 2010¹⁴).

Another rare process that is not simply affixation is reduplication. Reduplication is not widespread or common in this language family. One example is found in Labrador Inuttitut (Smith 1978b), specifically the Nain subdialect, where the intensifier *-pâ-* can be repeated. Speakers of related dialects remark on this property. A somewhat extreme elicited example is shown in (10).

- (10) Kuatsâ-pâ-pâ-pâ-pâ-pâ-pâ-lauk-Kunga
 startle-very-very-very-very-very-d.PAST-INDIC.1S
 ‘I really really had a very big fright’

DeReuse (1994: 90–1) also gives two examples from Siberian Yupik where productive doubling for emphasis is found.

A final rare occurrence are instances where the forms themselves are not phonologically conditioned, but their insertion is, that is, phonologically conditioned suppletion (see Carstairs 1990). An Uummarmiut (W) example from Lowe (1984: 157) is shown in (11) with two alternants *-pka-/ -tit-* ‘cause.’

- (11) a. niri-pka-raa
 eat-cause-DECL.3S/3S
 ‘She fed him’
 b. hiñik-tit-kaa
 sleep-cause-DECL.3S/3S
 ‘She put him to sleep’

Here the choice of causative allomorph depends on whether or not the form follows a stem ending in a vowel, leading to *-pkaq-*, or whether it follows a consonant, leading to *-tit-*. See Briggs et al. (in progress) for a dialect where the cognate morpheme allows both allomorphs only after vowels.

37.2 DERIVATION IN THE INUIT LANGUAGE

Having looked at the language family and the general processes involved in derivation, we now turn to a closer examination of the Inuit language with respect to types of

¹⁴ For example De Belder et al. (2010) give the non-compositional Italian *cas-ino* house-dim. ‘brothel’ for which they posit the diminutive to be generated below the category phrase nP and just above the root, while the compositional *nas-ino* nose-dim. ‘small nose’ would have the diminutive generated above the nP within an extended projection sizeP.

derivation. There are numerous dialects within this language and the discussion can be assumed to apply to all the dialects, although there is still much research to be done, as many dialects have received little study.

As noted in section 37.1, the set of derivational affixes in the Inuit language allows for a wide range of polysynthetic words. Fortescue (1980) calls the interaction of these postbases “internal syntax” and provides rewrite rules to generate them (see also de Reuse 1994). In particular, the nature of their interaction allows recursion, as shown by the Kalaallisut example in (12) from Fortescue (1984: 316).

- (12) aamaruti-ssar-siur- vi- ssar-siur- tu- tua-a-suq
 coal- FUT.-look.for- place FUT.-look.for- INTR.PART. only-be-INTR.PART.
 ‘who is the only one looking for a place to look/prospect for coal’

In this example we see that the verbal element *-siur-* ‘look for’ on the left can be within a complement of another higher *-siur-*, further on the right. This process appears to be limited only by processing.

That scope goes right to left is clearly linked to the inverse ordering of the morphemes. Higher syntactic elements occur towards the end of the word and lower ones closer towards the root. The meaning of a complex word can be best understood by reading the morpheme glosses right to left.

Some recent analyses treat these complex words as syntactically derived objects. Compton and Pittman (2010; see also Johns 2007, Cook and Johns 2009) take the syntactic nature of the postbases to be fundamental, and working within the framework of Distributed Morphology¹⁵ (Halle and Marantz 1993, etc.), provide an account whereby each postbase is generated in the syntax proper. Compton and Pittman argue that the grouping of elements into words results from a language-specific property in Inuit that requires that major syntactic domains (phases—see Chomsky 2008) be phonologically realized (or spelled-out) as words, that is, single uninterrupted sound units. If Compton and Pittman are correct, then the affixal nature of Inuit postbases is not a property inherent to each lexical item itself, but a predictable by-product of where the element is generated in the syntax. If it falls within the DP (noun-phrase) or CP (clausal) domain, it will form part of that complex unit. For Compton and Pittman, Inuktitut verbs are really clauses.

While the syntactic nature of postbases is generally accepted, there is nonetheless no principled way of distinguishing among these morphemes between those that are familiar as derivation in other languages, that is, involving category change, etc. and those that are usually thought to be too grammatical to be derivation, for example tense. The set of postbases is usually divided into four groups (i) verb modifying, (ii) noun modifying, (iii) verb making, and (iv) nominal making.¹⁶ I will divide these into two main

¹⁵ Under Distributed Morphology’s view of morphology (syntax all the way down—Harley and Noyer 1999), the distinction between word internal derivation and syntax does not exist.

¹⁶ This is the simplest way to present an overview. In fact, as Fortescue (1980) shows, there are classifications within these groupings, e.g. within verb modifying, some add higher structure (verb extending) while others simply modify.

categories (Woodbury 1981), (i) Change, which can involve not only change of category but change of valency, etc. and (ii) Modifying.

37.2.1 Derivation as Change

Category change is one of the more obvious types of change. Certain postbases change the category of the stem to which they are attached. Verbal stems can become nominal and nominal stems can become verbal.

37.2.1.1 *Verbs to Nouns*

The number of postbases that turn verbs into nominals is significantly fewer than the postbases that turn nominals into verbs. Some examples from the Kangiryuarmiut (W) dialect are in (13), taken from Lowe (1983). Beside each example, I give the semantic role of the derived nominal.

(13)	a.	uiguuq-ti-t	agent
		translate-AGENT-PL	
		‘interpreters’	
	b.	uuktu-un	instrument
		measure-instrument	
		‘ruler’	
	c.	hinig-vik	place
		sleep-place	
		‘bed’	
	d.	hava-qati	companion
		work-companion	
		‘co-worker’	
	e.	inuu-niq	event nominalizer
		alive-event.nom	
		‘life’	

Verbs can also be nominalized by the participial mood morphemes. In (14) we see Utkuhiksalingmiut (W) examples from Briggs et al. (in progress).

(14)	a.	hana-řaq ¹⁷
		work, make-PASS PART
		‘something that has been made’

¹⁷ This is arguably the same morpheme which is also found as an inflectional mood, as in the standard transitive participial mood forms (see Johns 1992, and see Fortescue 1995 for a historical account).

- b. hana-řuq
work-PART (one who)
'He's working (*verb*); one who works, a worker (e.g. store clerk or white-collar worker) (*noun*)

The example in (14a) is referred to either as the passive or transitive participle (Fortescue et al. 2010). The example in (14b) is referred to as the intransitive participle and is ambiguous between the nominal and the declarative in this dialect. Even though the participial can derive a nominal, it is often left out of postbase dictionaries and instead found in grammars. Nevertheless it clearly can nominalize, especially deep within the word, as in the Utkuhiksalingmiut (W) examples in (15).

- (15) a. apiri-řa-ksa-qaq-tunga
ask question-PASS PART-POT X-have-PART.1S
'I've got someone I can ask (*literally*: I have **someone** who can potentially be asked)'
b. qai-řu-qaq-tuq
come-INTR. PART-there is-PART.3S
'There is **someone** coming.'

In (15a) we see the passive participle to the immediate right of the root. This is followed by the postbase *-ksaq-* 'potential X,' a modifier which can only attach to nominals. Likewise the next morpheme to the right is the postbase *-qaq-* 'have/There is an X,' which attaches only to nominal stems, turning them into verbs (see section 37.2.1.2). Were the passive participle (or another nominalizer) not present, neither of these two postbases would be permitted.

In (15b) we see the intransitive participle, again immediately following the root. This in turn is followed by same *-qaq-* postbase as in (15a), which can only follow nominals and derives a verb. The nominalizing nature of the intransitive participial is particularly interesting because it sometimes seems to contribute only nominalization, followed almost immediately by category change back to verb, as in (16).

- (16) a. taqa-ju-mmari-alu-u-junga S. Baffin (E) (Johns 2010)
fatigue-PART-genuine-a.lot-be-PART.1S
 V N V
'I'm really tired'
b. taku-su-u-gamiuk Kalaallisut (Fortescue 1984: 285)
see-PARTICIPIAL-be-CAUS.3S/3S
 N V
'because HE saw it/ was the (only) one who saw it'

In the example in (16a), the intransitive participial *-juq-* nominalizes the verb, and is followed by a nominal modifier, *-mmarik* 'genuine' and then an intensifier *-aluk* 'a lot'. The construction subsequently reverts back to a verb when the copula *-u-* attaches. In (16b),

the verb root is followed by the intransitive nominalizer *-suq-* and then by the copula; the latter two form a common sequence (Fortescue 1984: 285).¹⁸

37.2.1.2 *Nouns to Verbs*

The Eskimo-Aleut language family is well-known for the property of adding postbases that change nominal stems to verb stems through what has been called noun incorporation (Sadock 1980, 1986, 2003). The Inuit language has quite a number of these postbases and recent discussion has concerned the semantics of the construction and the postbases as a set (Van Geenhoven 1998a, 1998b, Johns 2007). Some examples as shown in (17) from Johns (2007)

(17) a.	pitsi-tu-vunga dried.fish-consume-INTRANS.INDIC.1S. 'I'm eating dried fish.'	Labrador (E)
b.	tiki-lluq-tunga index.finger-have.a.sore-INTR.PART.1S 'I have a sore index finger.'	S. Baffin (E)
c.	qukiuti-taaq-tunga rifle-get-INTR.PART.1S. 'I got a rifle.'	Mittimatalik (E)
d.	inu-u-řunga inuk-be-INTR.PART.1S. 'I am an Inuk (Eskimo)'	Utkuhiksalingmiut (W)
e.	tipatsauti-sunniq-tuq perfume-smell-INTR.PART.3S 'It smells like perfume.'	S. Baffin (E)

¹⁸ Nominalization with the intransitive participial morpheme is also required in the derivation from roots to independent words which function in a similar way to adjectives and adverbs, as in the Baffin (E) examples (i) and (ii).

i.	anguti taki-juq taku-qqau-juq arnar-mit nait-tu-mit man tall-INTR.PART. see-T.PAST-INTR.PART.3S woman-mod. short-INTR.PART.-mod. 'The tall man saw the short woman.' Compton (2012: 87)
ii.	sukait-tu-mik ullar-tuq slowly-intr.part.-mod. run-intr.part.3s 'She runs slowly' Spalding (1993: 83)

In (i) we see two free-standing adjectives derived through the intransitive participial followed by case appropriate to their argument positions.

In (ii) we see that a free-standing adverb is derived through the addition of the intransitive participial plus default modalis case.

Such postbases have also been termed denominal (Mithun 1986). The denominal label refers to the fact that they obligatorily must attach to a nominal, even if it is the dummy root *pi-*, as in (18).¹⁹

- (18) *pi-si-juq* Baffin (E) Spalding (1998)
 dummy-find-INTR.PART.3S
 'He finds or comes across something'

The verbs realized by these postbases are analyzed by Johns (2007) as light verbs, to capture the fact that they are different from root verbs both in their requirement that a nominal root be attached, and by the fact that their meanings are restricted in range to that of very simple verbs, such as those meaning 'be,' 'have,' and elaborations on these meanings. Many of this set of verbs are antonyms or aspectual variants of a more basic one, for example *-qaq-* 'have' *-taa-* 'get.' No noun-incorporating verb conveys causal meaning or manner; thus verb meanings such as 'to tickle' or 'to clean' are not possible incorporating verbs.²⁰ A subset have evidential properties, as in (17e), expressing similarity in appearance, sound, and smell. Even though these verbs are often translated by English verbs with richer meaning, for example 'eat,' as in (17a), a closer examination shows that they are more abstract (see also Mithun 1999). Thus *-tuq-* often translated as 'eat' is more generally 'use, consume (food or drink) or wear X.'

In the majority of cases, the nominal stem²¹ to which incorporating verb attaches must be uninflected, as shown by the S. Baffin examples in (19a) and (19b), but a subset of incorporating verbs involving direction²² permit inflection, as in the Utkuhiksalingmiut (W) example in (19c).

- (19) a. *qimmi-qaq-tunga*
 dog-3S.POSS.-have-PART.1S
 'I have a dog.'
 b. **qimmi-nga-qaq-tunga*
 dog-3S.POSS.-have-PART.1S
 'I have his/her dog.'
 c. *qimatu-m-nungaq-tara*
 cached items-1S.POSS.-move to-PART.1S/3S
 'I put it in my cache of left-behind things.'

¹⁹ Under certain conditions the root can be elided (see discussion in Compton 2012).

²⁰ Yup'ik has verbalizing postbases which are variations on 'strike' (Woodbury 1981: 395), which can only attach to body parts.

²¹ Johns (2007) describes the incorporated element simply as a root but Compton (2012) shows that nominals with postbase modifiers are also found incorporated. This property is general across incorporation in the language.

²² See Sadock (1980); de Reuse (1994: 170) reports that this class exists in Siberian Yupik as well.

The nominals which attach to these verbal affixes often are referential (Sadock 1980). As shown in (20), a subsequent sentence can have the nominal as an antecedent.

- (20) Suulut timmisartui-lior-poq
 Søren plane-make-INTR.INDIC.3S.
 ‘Søren made an airplane.’
 Suluusa-qar-poq aquute-qar-llu-nii-lu
 wing-have-INTR.INDIC.3S. rudder-have-INF.-3RS-and
 ‘It has wings and a rudder.’ (Sadock 1980: 311)

The ability to have referential incorporating nominals varies across languages with noun incorporation. A further distinguishing feature is that incorporated nominals can even be names, or WH-elements, as in (21).

- (21) a. Fredi-jojâ-vutit uKauti-niat-tagâ
 Freddy-look.like-INDIC.2S. tell-N.FUT.-PART1S/3S
 ‘You look like Freddy_i. I am going to tell him_i.’ Labrador (E)
 b. kina-u-vit
 who-be-INTERR.2S
 ‘Who are you?’ S. Baffin (E)

Van Geenhoven (1998a) and Chung and Ladusaw (2003) argue that noun incorporation crucially involves nominals as semantic properties, and not entities, within a complex predicate. Under such a view, it should be impossible for names, which are inherently entities, to form part of such complex predicates; in other words names are predicted not to incorporate. While this is likely to be true of many languages with noun incorporation, examples such as (21a) demonstrate that the Inuit language requires a different analysis.

Generally the verbalizing postbases are followed by intransitive verbal agreement, although transitive²³ agreement can appear as well (see Johns 2009). The noun incorporating postbase *-gi-* ‘have as; consider’ is distinct in this property. It is obligatorily followed by transitive agreement, as in the Utkuhsalingmiut (W) example in (22).

- (22) ataata-gi-řara
 father-have.as-PART.1S/3S
 ‘He is my father.’

A second property that makes *-gi-* worthy of note is that it can combine with the nominalizing postbase *-vik* (see (13c)) to create a new argument for the verbal complex. The new argument is realized as an absolutive NP if overt, as in the Kangiryuarmit

²³ Agreement is labeled as intransitive vs. transitive but intransitive verbs can also take objects (see discussion of antipassive in Section 37.2.1.3.1). Better terms are single person vs. double person (Dorais 2010: 77), where single person refers to agreement with only one argument, and double to two arguments. Glossing within this article varies, depending on source.

(W) example in (23a). Both this example and the Baffin (E) in (23b) show transitive agreement.²⁴

- (23) a. qikiqtaq tangmaar-vi-gi-yaa Lowe (1983)
 island(abs) camp-place-have.as-TR.PART_{3S/3S}
 ‘He camped on the island’
 c.f. tangmaaqtuq ‘He camped’
 b. titirar-vi-gi-qattar-paatigut Spalding (1998)
 write-place-have.as-often-TR.INDIC_{3S/1P}
 ‘He regularly writes to us in Eskimo or Inuit language or style’
 c.f. titirartuq ‘He is writing’

This sequence, as mentioned in section 37.1.1, is considered to be lexicalized by Fortescue (1980), who analyzes it as a transitive postbase.

Another morpheme that stands out in this group is the postbase *-lik* ‘one who has X’ which appears to have properties both of a verbalizing and a nominalizing postbase, as in the Kangiryuarmiut (W) example in from Lowe (1983) in (24).

- (24) ilgaa-lik
 eyeglasses-one.with
 ‘the one with eyeglasses’

This postbase seems almost to be a portmanteau derivational morpheme in that it combines verbal (having/possessing) and nominalizing within the one form.²⁵

37.2.1.3 *Change in Grammatical Function*

There are a number of postbases that alter the valency of the verb root, either through permutation of how the arguments of the verb are realized syntactically, or the addition of arguments that are not part of the initial verb root, as we saw in (23).

37.2.1.3.1 *Permutation of Argument Realization*

Passive morphology affects the realization of arguments. The passive morphology in the Inuit language consists of the passive participle (see in (14a)) followed by the copula *-u-*, as shown in the Utkuiksalingmiut examples in (25).²⁶

²⁴ As Conor Cook pointed out to me (p.c.) this type of construction is interesting because the language has oblique cases, yet the postbase strategy prevails in these instances.

²⁵ Forms with this postbase not only appear as nominals but are deemed to be more appropriate as predicates than the competing verbal sequence *N-qaq-tuq* have-intr.part.3s ‘He/she has an X,’ but only in third person, e.g. *aKumigo-lik* ‘He has a speedboat’ Jeddore (1976) c.f. *aKumigok* ‘speedboat.’

²⁶ The agent can be optionally realized in either ablative (‘from’) or allative (‘to’) case, depending on the dialect; western dialects usually use the former, and eastern dialects the latter (although Kalaallisut uses the former—Fortescue 1984: 53).

- (25) kata-ga-u-řuq
 drop-passive-be-PART.3S
 ‘It was dropped (by somebody).’
 c.f. katak-tuq ‘it fell (by itself)’; katak-taa ‘he dropped it’

Another postbase which determines an alternate structure is the antipassive (also called semi- or half-transitive) morpheme (see Bittner 1987, Manning 1996, Spreng 2012). Both Bittner and and Spreng address the issue that the antipassive morphemes resemble aspect morphemes. Some examples of the antipassive morphemes from the Mittimatalik (E) dialect are shown in (26).

- (26) a. Peter kapi-si-vuq
 Peter (ABS) stab-AP-IND.3SG
 ‘Peter is stabbing someone.’ (Spreng 2012: 41)
 b. kunik-saq-tunga (Piita-mik)
 kiss-AP-PART.1SG (Peter-mik)
 ‘I am kissing (Peter)/someone (repeatedly, many kisses)’ (Spreng 2012: 29)

As we can see in (26a, 26b) the antipassive construction differs from a “transitive” in that there is single agreement with only the agent, not double²⁷ agreement (see footnote 23 on transitive agreement). The object of the action, if overt, is found in a case variously labeled as modalis, secondary case, or comitative. We see also that there are at least two forms of antipassive morpheme. Both forms are seen in other constructions as aspect markers, although sometimes with different phonological effects on the stem (see Spreng 2012).

Other postbases add arguments in addition to those of the stem. Two postbases which have this property are “order/ask/want” and “cause/let.” Examples are from the Labrador (E) dialect taken from Smith (1978a).

- (27) a. tiki-kqu-vauk
 arrive-ask/want-TR.INDIC.3S/3S
 ‘He wants him to arrive’
 b. taku-ja-u-ti-vauk
 see-PASS.PART.-be-cause-TR.INDIC.3S/3S
 ‘He shows it to him’

There is also a negative version of this postbase, for example the Kangiryuarmitut (W) -*ttaili* ‘to stop someone from X-ing’ (Lowe 1983). Pittman (2006, 2009) analyzes verbs such as those in (27) as restructuring verbs.

²⁷ Yupik can have transitive agreement following the antipassive, producing a malformative interpretation (Fortescue et al. 2010).

Another postbase which adds an extra argument is the benefactive or applicative morpheme *-uti-* ‘to do with or for’ (likely related to the instrumental—see section 37.2.1.1). This morpheme adds an argument for someone who either benefits, as in (28a), or reflexively undergoes the action, as in (28b), both examples are from Utkuhiksalingmiut (W).

- (28) a. niuv-uti-řa'ma
buy-do for-PART.2S/1S
'You bought something for me.'
- b. ukk-uti-řunga
close.door-do to oneself-PART.1S
'I closed the door on myself (and, e.g., crushed my fingers)'

Another class of postbases which adds on new arguments are those meaning “say” or “think” which take an entire proposition (TP—Pittman 2006) as a complement.

- (29) a. Miali igla-niraq-tara
Mary_(ABS) laugh-say-INDIC.1SG.3SG.
'I said that Mary laughed.' Baffin (E) (Pittman 2006)
- b. Jaani-up kapi-laung-niraq-taa tuktu Miali-mu
John-erg stab-dist.past-say-INDIC.3SG.3SG caribou(abs) Mary-mut
'(Last week) John said that Mary stabbed the caribou.' Baffin (E) (Pittman 2006)

As Fortescue (1984), Woodbury (1981), and Sadock (1986) point out, this is the only class of verb-like postbase which can follow tense morphemes, as seen in (29b).

37.2.2 Modification

The next major class of postbases are those which change neither the category or grammatical function properties of the stem to which they attach, but instead add extra meaning to it.

37.2.2.1 Noun Modifying

Postbases which attach to the nominal without changing category are similar to what is termed evaluative morphology (diminutives, augmentatives, etc.). The members of the class are typical in meaning for evaluatives, and include size, affective terms, age, quality, etc.; less typical are ones meaning former, as in (30b), fake/toy, as in (30c), genuine, potential, as in (15a). The majority have both positive and negative values.

The set that modifies nominals is again smaller than those which modify verbal stems. Some examples from S. Baffin (E) and taken from Compton (2012) are shown in (30).

- (30) a. umingma-jjuaq
 muskox-big
 'the/a big muskox'
 b. iglu-viniq
 house-old/former
 'the/an old house'
 c. nanu-ralaa-nngua-t
 polar.bear-small-pretend-PL.
 'small pretend polar bears'
 d. anaana-tsia-kuluk
 mother-good-adorable
 'dear/adorable grandmother'

As (30c, 30d) show, it is possible to combine, or stack, multiple members of this class in one word. Compton (2012) also reports that a re-ordering will not necessarily change meaning. Where a postbase forms part of a lexicalized combination with a root, as in (30d), where *anaanatsiaq* means 'grandmother,' ordering is important (see also Mithun 1999 on Yupik).

Compton argues that this class of morphemes consists of strictly-attributive adjectives, thus delimiting the set from all possible adjectives. In other words, adjectives that have intersective properties cannot be members of this set of postbases (and indeed are found as independent words rather than affixes). The core idea is that if the meaning of an adjective cannot exist independently of a root, it will be a postbase in the Inuit language. Thus one can say of a red house that it is both a red thing and a house thing, but with strictly-attributive adjectives, as in 'good house' one cannot say that there are good things and also a house things. Its goodness is directly dependent on how good a house it is.

There are also a few non-evaluative modifiers which can attach to nominal stems. One example is a postbase in Utkuhiksalingmiut (W) meaning "only":

- (31) iniqni-inna-it
 adult-only
 only adults (no children)

This postbase also modifies verbal stems; there is a division within modificational postbases as to whether or not they are restricted to a particular category.

Some noun modifying postbases appear to play a larger role than just modifier. For example, *tsa(k)* in the Labrador dialect (E) is translated in Smith (1978a: 91) as 'something for x; a future x, a potential x.' This basic meaning is shown in (32).

- (32) sini-tsak Jeddore (1976)
 border-POT
 'binding material or other material used to make border on a coat or a silapāk
 (outer shell of parka)

Here the noun plus modifier together refer to a substance or material that has the potential to become the type of entity indicated by the root.

We see the same postbase providing the meaning equivalent to English preposition *for* or *to* as in the following typical Labrador (E) example.

- (33) a. allatigi-uti-tsa-Ka-ven
 wipe-instrument-POT-have-INTERR.2S
 ‘Do you have something to wipe it with?’ (Jeddore 1976)
- b. kiasaleni-tsa-vuttinik
 gasoline-POTENTIAL-POSS.1PL.MODALIS
 ‘for our gasoline’
 Meaning context: ... ‘and if we have enough money for our grub and *for our gasoline*.’ (Winters 2007)

The meanings here are slightly different from (32). Example (33a) does not refer to something that will literally be turned into a cloth but instead something which will be used as a cloth. In (33b) the context of the utterance is within a discussion of having enough money ‘for our gasoline.’ Here again, the use is more abstract in that rather than indicate something which could become gasoline, it instead indicates that the gasoline is a hypothetical entity (although likely). This gives approximately the equivalent in meaning to the English preposition *for* X. As a result this postbase is used where English uses *to* as in (33a), which discusses a hypothetical event, and *for* in (33b), which refers to a hypothetical entity. Arguably these are more complex usages than (32).

Another noun modifying postbase -*vinik* is usually translated as ‘late’ or ‘former’ X, but again it is used more frequently than *former* is used in English, as can be seen in the Labrador (E) example in (34).

- (34) pullauja-vinik
 bottle-former
 ‘a broken bottle’ (Jeddore 1976)

While English speakers could probably understand *former bottle* to mean one which is broken, it is not a natural way in English to express this meaning.

37.2.2.2 Verb Modifying

The final class of postbases we will examine are those that modify verbs. This is a very large class, and perhaps the one that seems the less familiar as derivational morphology. Nonetheless, as mentioned previously, there is no principled reason to single out this class from the other postbases as being essentially different. Given the clausal nature of the verbal complex, it is unsurprising that verb modifying postbases consist of aspectual markers, all non-derived adverbs (Compton 2012), tense, negation, verbs taking same subject verbal complements (e.g. “want” which Johns (1999) analyzes as a modal), and even postbases conveying attitudes of the speaker. Readers are referred to Fortescue

(1980), Woodbury (1981), de Reuse (1994), and Compton (2012) for more considered and detailed discussion of this class.

We will briefly look at some sub-groups within this class. One group is negation, which takes two forms, one which follows stative or adjectival stems, as in (35a), and the more general negative, as in (35b), both examples from Utkuhiksalingmiut (W).

- (35) a. naku-it-tuq
strong-not-PART.3S
'He is weak.'
b. taku-ngngi-taa
see-not-PART.3S/3S
'He doesn't see it.'

Another sub-group are tense and aspect morphemes. Not all dialects require obligatory past tense marking, for example Kalallisuut (Bittner 2005), and some lack past tense morphology altogether, for example Utkuhiksalingmiut (Jean Briggs p.c.). Hayashi (2011) argues that in at least some dialects, there are past tense postbases, and that they represent hodiernal (restricted to today) and non-hodiernal tenses. Some examples of past usage are shown in (36).

- (36) a. ii, tusaq-vigi-qqau-jara ullumi
yes hear-from-R.PAST-TR.PART.1S/3S today
'Yes, I heard from her today' Baffin (E) (Hayashi 2011: 50)
- b. qanga tiki-lauq-paa
when arrive-d.PAST-INTERR.3S
'When did he arrive?' Baffin (E) (Hayashi 2011: 49)

Yet another sub-group within this large class might be described as adverbials, which Compton (2012) shows can be stacked.

- (37) a. ani-saaq-tuq
go.out-quickly-INTR.PART.3S.
'He/she left quickly/just left.' S. Baffin (E) (Compton 2012: 1)
- b. ani-kasa-kkanni-ngaaq-tuq
go.out-almost-again-instead-INTR.
PART.3S
'He/she almost left instead again.' S. Baffin (E) (Compton 2012: 133)

37.3 CONCLUSION

We have seen that the Eskimo-Aleut language family exhibits a strikingly syntactic instantiation of derivational morphology, where the morphology within the word extends, permutes, and modifies the initial root. The productivity, complexity, and general transparency of the relations of the pieces within these words leads us to analyze them as clauses. Nevertheless derivational properties within these words are also prominent. The right to left ordering of their components conforms to derivational orderings observed in other languages. In addition, we saw that category change plays a fundamental role within these word-clauses. Finally, some of the more conventionally derivational morphemes, such *-tsa(k)-* ‘potential X’ or *-vik* ‘place for X-ing’ were seen to also be found in contexts where their basic properties were harnessed to express complex syntactic structure. What we see in this language family is a larger distribution of word-internal mechanisms of category change and modification than is seen in more familiar languages. Rather than being restricted to relatively simple structures, these morphemes play a central role throughout the grammar.

CHAPTER 38

UTO-AZTECAN

GABRIELA CABALLERO

38.1 INTRODUCTION

THIS chapter addresses derivation mechanisms in the Uto-Aztecan language family through the lens of Rarámuri (Tarahumara), a Uto-Aztecan language of Northern Mexico. Through examination of representative derivational categories and processes in Rarámuri, this chapter situates derivation within the larger scheme of the agglutinating morphology of this language. Some of the more general theoretical and typological issues that the descriptive facts raise concern the nature of recursivity of derivational morphology, proposed general principles of affix order, as well as the intricacies of the morphophonological processes involved in marking these processes.

This chapter is structured as follows. Section 38.2 provides a brief overview on the Uto-Aztecan language family. In Section 38.3, I characterize recurrent aspects of Uto-Aztecan morphological systems and describe three derivation strategies in Rarámuri, which are well-represented in Uto-Aztecan. This section focuses on both the categories overtly marked, as well as the morphophonological processes associated with both the concatenative and non-concatenative constructions deployed to mark these constructions. Section 38.4 concludes with a summary.

38.2 THE UTO-AZTECAN LANGUAGE FAMILY

The Uto-Aztecan (UA) language family is the largest language family of the Americas in terms of number of languages, with more than sixty varieties still spoken by large numbers of speakers. The vast territory where UA languages were and are still spoken ranges from the Great Basin area in North America (with languages from the Numic branch) to El Salvador in Central America (Pipil (Aztecán)). Figure 38.1, based on classifications

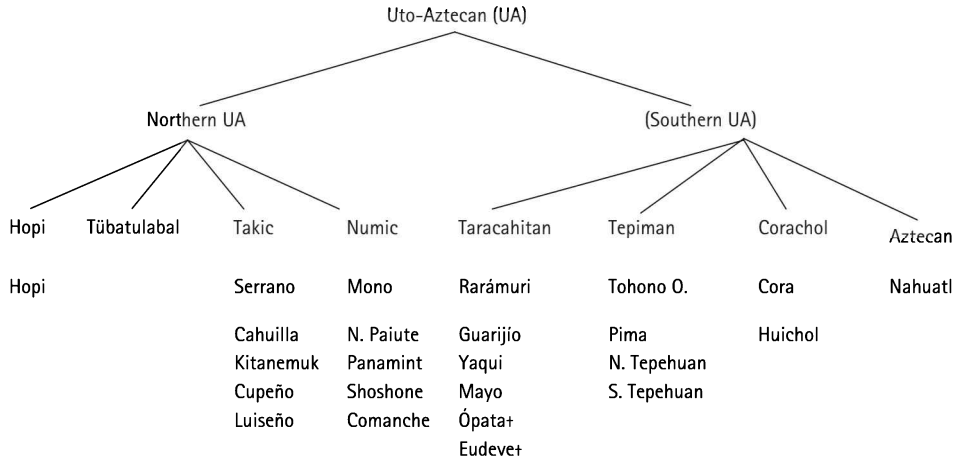


FIGURE 38.1 Uto-Aztec language family

Notes: Individual branches are shown in two main groups, Northern UA and Southern UA, a grouping that largely matches the geographical border between the USA and Mexico. Northern UA has solidly been established as a genetic unit (Heath 1978, Manaster Ramer 1992), though evidence for the same status for the Southern branch is still debated (Kaufman 1974a, 1974b, Campbell and Langacker 1978, Heath 1978, Miller 1983, 1984, Cortina Borja and Valiñas Coalla 1989, Hill 2001). Languages that have no speakers left are marked with†.

from Langacker (1977), Campbell (1997), and Mithun (1999), provides a simplified schema of individual languages and their genetic affiliation into major subgroupings.

Research on UA languages has placed a great deal of attention on the complex morphological systems these languages display.¹ Despite the great diversity they possess—a byproduct of their large geographical extension and time-depth (estimated to range between four and five thousand years (Fowler 1983, Campbell 1997, Silver and Miller 1997, Hill 2010)), UA languages also display a number of recurrent typological patterns and structures. In terms of their morphologies, these recurrent patterns include head-marking, complex prosodic morphology patterns, and large derivational paradigms (Sapir 1921, Whorf 1935, Haugen 2008b).

38.3 DERIVATION IN UTO-AZTECAN: RARÁMURI

Before addressing derivation categories and processes in UA, it is first necessary to characterize the nature of the morphological systems of these languages. One of the most relevant properties of these systems is the fact that they display canonical agglutinating

¹ Detailed descriptions of the morphological systems of individual languages can be found in Sapir 1930 (Southern Paiute), Saxton 1963 (Tohono O'odham), Hale 1965 (Tohono O'odham), and McLaughlin 1987 (Panamint), among others.

properties. Agglutination has been defined through the conflation of the parameters of fusion and flexion, such that agglutinating morphologies are systems that are concatenative and nonflexive. While the notion of agglutination in morphological typology has been questioned as being largely inconsequential as a defining characteristic of morphological systems (Spencer 1991, Haspelmath 2009),² it has also been proposed that splits in the expression of properties along an agglutination/fusion cline are systematic and meaningful at a large scale (Plank 1999). Under the latter view, a finer characterization of agglutination involves making reference to a large set of parameters, which includes the presence/absence of zero exponence, homonymous exponence, optional marking, large paradigms, and the degree of phonological transparency in formal marking, among others.

UA languages possess fairly elaborate verbal morphology, and display most of the agglutinating-like properties enumerated above, including a low degree of phonological cohesion between exponents, and a low degree of cumulation in morphological exponence (Langacker 1977: 158). UA languages also display rich word formation systems with varying degrees of productivity. These systems feature deverbal nominalization (including agentive, patient, and event nominalizations (e.g. Numic (Toosarvandani 2010)), compounding, conversion, denominal verb formation, and rich argument structure changing derivational systems.

In this section I do not attempt to give an exhaustive characterization of derivational constructions in UA, but rather provide a detailed overview of three derivation categories and processes that are well represented across UA: (i) instrumental prefixes, (ii) denominal verbalizing morphology, and (iii) argument structure changing morphology. In order to characterize these categories and situate them within the larger context of the morphological system of a particular language, I provide the details of these individual processes in Rarámuri, a UA language of the Taracahitan branch.³

Rarámuri is spoken by approximately 95,000 speakers (Lewis 2009) in the Mexican state of Chihuahua. There is a high degree of variation among different Rarámuri varieties, which comprise a dialect continuum.⁴ The data presented in this chapter correspond to the variety spoken in Choguaita, spoken by approximately 1,000 speakers (Casasus in prep.). Despite being spoken by the community at large, growing marginalization

² “[L]anguages as wholes tend to involve greater or lesser mixtures of different sorts of processes. . . [and] it is individual rules and not entire languages that can be meaningfully called agglutinating” (Anderson 1992: 328–9).

³ Examples are presented in IPA, with a couple of exceptions: the phoneme /r/ and its word-initial allomorph /r/ are represented as <r>; and the lateral flap, which has a retroflex-like articulation [ɽ] and a lateral-like articulation is represented as <l>. Each CR example gives information about the source of that particular example (in square brackets) indicating the contributor’s initials, year the example was recorded, book and page number (if source example is taken from fieldnotes) or electronic document identifier, and type of document where the data come from: elicitation (el), text (tx), interview (interv), etc. Some examples provide a free Spanish translation, the language used for elicitation.

⁴ The *Ethnologue* (Lewis 2009) proposes five major dialect areas of Rarámuri, but the amount and nature of variation among these language varieties is still largely understudied.

and political instability in the area have intensified the degree of endangerment of the language.

Choguita Rarámuri (henceforth CR) morphology is almost exclusively suffixing and also displays many of the canonical agglutinating properties present elsewhere in UA, including mostly concatenative exponence, potentially long string of suffixes, virtually no flexivity, zero exponence, little homonymous exponence, multiple exponence, and abundant optional marking. While CR derivation is generally realized through concatenative markers, stem allomorphy and several other non-concatenative processes are also deployed in marking this kind of information.

38.3.1 Instrumental Prefixes

A well-known phenomenon of UA languages is the presence of a set of instrumental prefixes that indicate the instrument with which a transitive activity is done, or the manner in which the activity is carried out. Dayley (1989: 92) reconstructs these prefixes (approximately 20) for Proto-UA. A list of representative examples is provided in (1).

(1) Uto-Aztecán instrumental prefixes

	<i>Instrumental Gloss</i>	<i>Reconstructed roots</i>	
		<i>Form</i>	<i>Gloss</i>
a.	'with heat or fire'	*kuh	'fire'
b.	'with the teeth or mouth'	*kü'i	'bite'
c.	'with the hand'	*maa	'hand'
d.	'with the nose'	*mu-pi	'nose'
e.	'(with or pertaining to) water'	*paa	'water'
f.	'with the butt or behind'	*pih	'back'
g.	'with or from cold'	*süp	'cold'
h.	'with the mind, by feeling or sensation'	*suuna	'heart'
i.	'with the foot'	*tannah	'foot'

Many of the instrumental prefixes derive from body part terms, but these formatives are not restricted to this word-class (e.g. (1a, e, g)). These formatives are synchronically active in the Numic and Tepiman branches, and have only lexicalized remnants in the rest of the UA family. This is true in the case of CR and closely related River Guarijío (Miller 1996), where instrumental prefixes are only found in a few lexicalized items, as exemplified in (2).

(2) Instrumental prefixes in CR and River Guarijío

<i>CR</i>			
a.	<i>ma-tŕō</i>	'hand-hit' 'to hit with the hand'	[RF 04 1:124/el]
b.	<i>ma-hó</i>	'hand-dig' 'to fondle'	[BFL 04 Com/el]

c. <i>rara-hó</i>	'foot-dig'	'to dig with the foot'	[SFH 08 1:48/el]
d. <i>sika-tfô</i>	'hand-hit'	'to hit with the hand'	[SFH 04 1:123/el]
<i>River Guarijío</i>			
e. <i>magopá</i> ⁵		'to carry with the hand'	[Miller 1996: 355]

In some cases, the instrumental prefix does not have a correlated nominal form that is synchronically productive, as the formative *ma-* in (2a, 2b). Lionnet (2001: 87) identifies several instrumental prefixes in Norogachi Rarámuri, including *ba(?)*- 'water,' *ku-* 'wood,' *ma-* 'hand,' *mo(?)*- 'head,' and *na-* 'fire,' none of which have a corresponding independent noun. While not all of the prefixed formatives have been lexicalized in CR (e.g. *rará* 'sole of foot' (2c), *siká* 'hand' (2d)), instrumental prefixation is no longer a productive process in this or any of the other languages of the Taracahitan branch.

Instrumental prefixation resembles body part incorporation, a restricted kind of noun incorporation that is common in the Uto-Aztecan family and other languages of the Americas (Baker 1996, Sapir 1911). Both instrumental prefixation and body part incorporation can be classified as subtypes of N-V compounding or Type I Noun Incorporation in Mithun's (1984) typology (Haugen 2008b: 118). In body part incorporation, a type of object incorporation,⁶ the incorporated nouns are restricted to referring to body parts and bodily fluids and they precede the verbal root, the head of the incorporated construction. This process is present in CR, as exemplified in (3).⁷

(3) *CR Body-part incorporation*

a. <i>moʔo-répi-ri</i>	'head-cut-PST'	'to behead, INTR'	[SFH 07 1:185/el]
b. <i>busi-kási</i>	'eye-break'	'to become blind'	[SFH 06 1:112/el]
c. <i>ropa-kási</i>	'belly-break'	'to have a miscarriage'	[BFL 07 VDB/el]
d. <i>siwa-bóti-ri</i>	'guts-loosen- PST'	'to disembowel, INTR'	[SFH 07 1:186/ el]
e. <i>kuta-bíri</i>	'neck-twist'	'to twist neck, INTR'	[SFH 07 VDB/el]
f. <i>tŕoma-bíwa</i>	'mucus-clean'	'to clean mucus'	[SFH 08 1:51/el]

Instrumental prefixes, body part incorporation and other unproductive compounding-like processes are part of an Inner Stem domain in the CR verbal complex. This verbal domain, the input to suffixation, is composed of a denominalized noun

⁵ This example respects Miller's original transcription, which provides this form without any internal analysis. It is possible to assume, however, that while no synchronically independent verb is listed in Miller's glossary, the example contains the *ma-* 'hand' prefix found in the Rarámuri examples and reconstructed for Proto-UA.

⁶ As Haugen (2008b: 119–20) notes, while instrumental prefixation overwhelmingly involves a meaning of a manner or instrument with which an activity is carried out, these forms can also be found with object readings, as in Comanche (Charney 1993).

⁷ Many of these constructions are fully compositional and have transparent semantics (e.g. (3a, 3d–3f)), but in many cases the forms have been lexicalized (e.g. (3b, 3c)). The incorporated noun generally fills the syntactic role of object, decreasing the verb's valence, and may be externally modified. Incorporated body part nouns, however, do not form an open class, since only a few body part terms are found in these constructions, and not the full range of lexical items in this semantic class.

root, an incorporated verb root, or a verbal root that has optionally undergone a non-concatenative process or a process that is no longer fully productive, including subject number prefixation, pluractional marking, and denominal suffixation and conversion (the latter is addressed in Section 38.3.2). Morphophonologically, the CR Inner Stem is characterized by being the domain of application of compensatory lengthening and other phonological processes. These processes make the morphological constructions taking place at this stage more tightly bound phonologically to the root than any later morphology.

38.3.2 Denominal Verbalizing Morphology

Denominal verbs are also robustly attested across UA languages. They typically involve meanings such as “make,” “have,” “put on,” “use,” “remove,” “get,” “marry,” and “become,” among others, and have similar morphosyntactic characteristics across the family (Haugen 2008a). CR possesses several denominal verbal suffixes of varying degrees of productivity, some of which I describe below. Within the CR verb, these suffixes are part of the Inner Stem domain defined in Section 38.3.1.

A noun that adds the suffix *-ta/-ra* derives a verb meaning “to make/become N” (4a) or, when used with nouns referring to a piece of clothing, “to wear N” (4b, 4c). In other Tarachitan languages, cognate forms of this suffix may also have a “use” meaning (e.g. Yaqui, Dedrick and Casad 1999: 138; cited in Haugen 2008a: 459). The derived form in CR requires stress to be in the denominalizing suffix. When this suffix is added to a trisyllabic noun, the last syllable of the noun is truncated in order to meet the phonological requirement of this construction (e.g. (4c)).

(4) Verbalizer *-ta/-ra* “make/become”

<i>Form</i>	<i>Gloss</i>	<i>Translation</i>	<i>Noun</i>	
a. <i>nori-rá-</i>	cloud-VBLZ-	‘to get cloudy’	<i>norí</i>	[BFL 04 1:92/el]
b. <i>aka-rá-</i>	sandal-VBLZ-	‘to wear footwear’	<i>a^hká</i>	[SFH 08 1:47/el]
c. <i>sipu-tá</i>	skirt-VBLZ	‘to put on a skirt’	<i>sipút/â</i>	[BFL 07 Sept 6/el]

The allomorphy of this suffix is part of a larger set of alternations in CR involving stops ([p ~ b], [t ~ r], and [k ~ g]). Many of these alternations show up synchronically as suppletive allomorphy that is lexically determined,⁸ as in the case of denominal *-ta/-ra*. This is also the case for another denominal verbalizing marker, the *-tu/-ru* ‘gather/bring’

⁸ These stop alternations have been argued to be a remnant of historical alternations involving voiceless/fortis stops and voiced/lenis stops, a phenomenon that is synchronically active in Numic (e.g. Southern Paiute (Sapir 1930), Kawaiisu (Zigmond et al. 1991)). Some of the stop alternations in CR can be attributed partially to phonological conditions, since there is a phonological process that devoices voiced/lenis consonants after another consonant (a derived environment stemming from stress-conditioned syncope) (Caballero 2008).

suffix, a marker of limited productivity that is always stressed in the derived verb.⁹ It attaches to the nominal root, and as part of the Inner Stem, it is sensitive to the morphophonological effects that characterize this stem level.

(5) Verbalizer -tu suffix

a.	<i>baʔwi-tú</i>	‘water-GATHER’	‘to bring water’	
b.	<i>ila-rú</i>	‘cactus-GATHER’	‘gather cactus’	[SFH 08 1:52/el]
c.	<i>naʔi-rú</i>	‘fire-GATHER’	‘get fire’	[SFH 08 1:52/el]
d.	<i>raki-rú</i>	‘palm-GATHER’	‘gather palm’	[SFH 08 1:52/el]

The suffix *-bu* is another suffix of limited productivity that derives intransitive verbs from nouns with the meaning “remove” or “undo” (e.g. (6a–6c)). A cognate suffix can be found in River Guarijío, with stress-based allomorphy (*-pu* when stressed, *-bu* when unstressed) (Miller 1996: 151) (e.g. (6d–6f)). In River Guarijío, this suffix can attach to inchoative stems (e.g. (6f)).¹⁰

(6) Reversive -bu suffix in CR and River Guarijío

CR

a.	<i>boʔo-bú-ma</i>	‘feather-REV-FUT.SG’	‘to pluck (a chicken)’	[SFH 08 1:51/el]
b.	<i>riʔi-bú-a</i>	‘stone-REV-PROG’	‘to take out stones’	[SFH 08 1:51/el]
c.	<i>tʃomo-bú-ka!</i>	‘mucus-REV-IMP.SG’	‘to clean nose’	[SFH 08 1:51/el]

River Guarijío

d.	<i>etʃapóa-bu-na</i>	‘beard-REV-PRS’	‘to shave the beard’	[Miller 1996: 151]
e.	<i>sawa-pú-na</i>	‘leaf-REV-PRS’	‘to remove the leaves’	[Miller 1996: 151]
f.	<i>kahka-pá-bu-na</i>	‘sweet-INCH-REV-PRS’	‘to become less sweet’	[Miller 1996: 151]

Finally, there is a denominal verbalizer *-e* suffix with the meaning ‘to have’ or ‘to wear’ that replaces the final stem vowel. Like the rest of the denominal suffixes described in this section, it is always stressed. The examples in (7) show this derivation.

(7) CR Verbalizer suffix -e

a.	<i>kun-é-</i>	‘marry a man’	<i>kuná</i>	‘husband’	[BFL 05 1:116/el]
b.	<i>up-é-</i>	‘marry a woman’	<i>upí</i>	‘wife’	[BFL 05 1:116/el]
c.	<i>ak-é</i>	‘wear sandals’	<i>aká</i>	‘sandal’	[BFL 05 1:116/el]
d.	<i>wir-é</i>	‘wear earrings’	<i>wirá</i>	‘earring’	[BFL 07 sept 6/el]
e.	<i>motos-é</i>	‘have white hair’	<i>motosá</i>	‘white hair’	[BFL 07 sept 6/el]

⁹ That is, while not attested with all nominal roots, this suffix still produces novel words.

¹⁰ This suffix has a corresponding, synchronically active verb: the verb *puha*, ‘to take away’ in River Guarijío (Miller 1996: 151) and its Rarámuri cognate form *bu(?)é* ‘to take away’.

In addition to these concatenative processes, CR has denominalizing constructions that involve stem alternations and conversion. Some nouns can be denominalized through a stress shift within the stem. CR possesses two kinds of suffixes in terms of their stress properties, namely stress-neutral suffixes, which cause no stress changes in the words they appear in, and stress-shifting suffixes, which condition stress shifts with certain roots (Caballero 2008, 2011). The stress shift in denominalizing constructions cannot be attributed to the effect of stress-shifting suffixes. Rather, the stress shift itself is the marker of the construction; in this case the derived verbs have third syllable stress, while the corresponding nouns have second syllable stress. This is exemplified with denominal verbs inflected with a stress-shifting suffix (8a) and a stress-neutral suffix (8c). The stress shift in both of these forms is attributable to the verbalizing process, not a phonological effect conditioned by a following suffix.¹¹

(8) Verbalizing stress shift

	<i>Form</i>	<i>Gloss</i>	<i>Stress position</i>	
a.	<i>siputŷā-ma</i>	‘wear.skirt-FUT.SG’	3 rd syllable	[BFL 07 Sept 6/el]
b.	<i>sipútŷa</i>	‘skirt’	2 nd syllable	[BFL 07 Sept 6/el]
c.	<i>napatŷā-ri</i>	‘wear.blouse-PST’	3 rd syllable	[BFL 07 Sept 6/el]
d.	<i>napátŷa</i>	‘blouse’	2 nd syllable	[BFL 07 Sept 6/el]

Finally, some nominal stems (including some nouns referring to weather) can take verbal morphology with no overt denominalizing marking, a case of conversion. The examples in (9) are listed with the future singular suffix. With some exceptions (like 9a), most of these verbs derived through conversion are incorporated as stems where stress shifts take place after the addition of stress-shifting suffixes.¹²

(9) Conversion

	<i>Verb</i>	<i>Gloss</i>	<i>Noun</i>	<i>Gloss</i>
a.	<i>remé-ma</i>	‘make.tortillas-FUT.SG’	<i>remé</i>	‘tortillas’
b.	<i>ika-méa</i>	‘be.windy-FUT.SG’	<i>iká</i>	‘wind’
c.	<i>moro-méa</i>	‘be.foggy-FUT.SG’	<i>morí</i>	‘smoke’
d.	<i>uku-méa</i>	‘rain-FUT.SG’	<i>ukí</i>	‘rain’
e.	<i>nori-méa</i>	‘be.cloudy-FUT.SG’	<i>norí</i>	‘cloud’
f.	<i>saki-méa</i>	‘make.esquite-FUT.SG’	<i>sakí</i>	‘esquite’

¹¹ Some of the nouns that undergo this stress shift may alternatively become denominalized through other processes. For instance, *siputŷā-ma* can be alternatively realized as in (4c), *sipu-tá-ma* ‘skirt-VBLZ-FUT.SG’ with denominalizing suffix -*tá* (and concomitant noun truncation). I have not detected any semantic difference between these alternative forms.

¹² Allomorphy of the future singular suffix in CR is stress-based (-*méa* is the stressed allomorph, and -*ma* the unstressed one).

Examples (9c, 9d) also show that the unstressed vowel of the derived verb may harmonize with the first vowel of the stem (e.g. *morí* ‘smoke’ becomes *moro-* ‘to be smoky/foggy’ and *uki* ‘rain’ becomes *uku-* ‘to rain’). This harmonizing process, where the root final high vowels agree in color with the preceding vowel once they are unstressed, is absent in cognate forms of zero derivation in Guarijío (cf. *yu’kí* ‘rain,’ *yu’ki-má* ‘to rain’) (Miller 1996: 148).

Outside of Taracahitan, denominal verbalizing morphology may display other functions, such as marking inalienable/alienable distinctions through distinct formatives marking possession (“have N”) (e.g. Numic (Haugen 2008a: 452)).¹³ Most notably, denominal verbs in many UA languages display properties attributed to noun incorporation, including modifier stranding and the presence of hyponymous object arguments (Haugen 2008b). In CR, denominalized verbs may take an object argument that refers to a subset of the entities denoted by the noun that undergoes denominalization. This is illustrated in (10).

(10) CR hyponymous object with denominal verb

nihé	aka-rá-sa	sapato
1SG.NOM	sandal-VBLZ-COND	shoes
‘I will wear shoes’		
‘Voy a ponerme zapatos’ [SFH 08 1:47/el]		

As Haugen (2008b) notes, this is a property that characterizes classificatory noun incorporation (Mithun 1984), and is attested elsewhere in denominal verb formation in UA (Hopi; Yaqui (Taracahitan); Tümpisa Shoshone, Gosiute Shoshone (Numic); and Cupeño (Takic) (Haugen 2008b: 445–6).

38.3.3 Argument Changing Morphology

Perhaps one of the most distinctive characteristics in UA derivation is the presence of exuberant argument structure changing morphological patterns. In CR, several valence-increasing constructions are available, depending on the phonological and semantic properties of verbal roots.

Verbal roots in CR can be characterized as either stressed or unstressed, with significant differences in their prosodic makeup, depending on their interaction with different types of suffixes (Caballero 2008, 2011). Unstressed roots undergo stress shifts when combining with stress-shifting suffixes. A subset of these unstressed roots may participate in a valence stem allomorphy system, undergoing a series of non-concatenative

¹³ In Yaqui, the distinction between alienable and inalienable possession in denominalization is carried out through affixation and a periphrastic construction, though the distinction between constructions is not consistently correlated with a difference in meaning (Jelinek and Escalante 1988: 419, cited in Haugen 2008b: 447–8).

changes when specified for intransitive, transitive, or applicative meanings. Another subset of the unstressed roots, semantically defined as a class of verbs of change of state, participates in a second valence stem allomorphy system characterized by the use of thematic suffixes. In both of these systems, final stem vowel changes and fixed stress patterns are the markers of morphological categories (valence-increasing categories), while any concomitant stress shifts are epiphenomenal of certain morphological constructions, a byproduct of the phonological makeup of roots and affixes and their compositional interaction.

38.3.3.1 Valence Stem Allomorphy

In the valence stem allomorphy system there are three kinds of stems: intransitive, transitive, and applicatives.¹⁴ Intransitive stems end in an unstressed vowel (11a, 11b), transitive stems end in a stressed, low mid vowel (11c),¹⁵ and applicative stems end in a stressed front vowel (11d–11f). In these cases, applicatives involve the introduction of a benefactive or a malefactive argument.

(11) Valence stem allomorphy

Intransitive

- a. *nihé* *ma* *noko-méa*
 1SG.NOM already move.INTR-FUT.SG
 ‘I will move already’
 ‘Ya me voy a mover’
 [SFH 05 1:80/el]
- b. *uku-méa*
 rain.INTR-FUT.SG
 ‘It will rain’
 ‘Va a llover’
 [SFH 05 1:82/el]

Transitive

- c. *ma* *rará-ki*
 already buy.TR-PST.1
 ‘(He) already bought it’
 ‘Ya lo compró’
 [AH 05 1:130/el]

Applicative

- d. *nihé* *mi* *troka* *noké-ri*
 1SG.NOM 2SG.ACC truck move.APPL-PST

¹⁴ While not all unstressed roots participate in this system, it cannot be merely understood as the residue of a diachronic development.

¹⁵ Heath (1978) argues that an intransitive–transitive contrast marked by *i* for intransitive and *a* for transitive goes back to Proto-Uto-Aztecan.

- ‘I will move the truck for you’
 ‘Te voy a mover la troca’
 [SFH 05 1:80/el]
- e. *á=mi tamí raré=ma*
 AFF=2SG. 1SG.ACC buy.APPL-FUT.SG
 NOM
 ‘Will you buy (it) from me?’
 ‘Me lo compras (yo lo vendo)?’
 [SFH 05 1:75/el]
- f. *mi uké=ri*
 DEM rain.APPL-PST
 ‘It rained on (him)’
 ‘Le llovió’
 [SFH 05 1:84/el]

As the examples in (11a, 11b) show, the final unstressed vowel of a stem may harmonize with the first stem vowel (the noun *ukí* ‘rain,’ has a final high front vowel, and the stressed alternant of *noko-méa* is *nokí*, with a final, stressed *i*). Table 38.1 schematizes the three-way contrast of these stems (syntactic/semantic gaps are symbolized by dashes).

Finally, a morphophonological property of these stems is that they block any stress shift or vocalic alternations expected to occur when combining with stress-shifting suffixes. That is, transitive and applicative stems always have stress in the final vowel, regardless of the phonological properties of the suffixes they combine with.

Table 38.1 Valence stem allomorphy

Intransitive	Transitive	Applicative	Gloss
a. <i>suwí</i>	<i>suwá</i>	<i>suwé</i>	‘run out/finish up’
b. <i>sawí</i>	—	<i>sawé</i>	‘cure, heal’
c. —	<i>rará</i>	<i>raré</i>	‘buy’
d. <i>noko</i>	—	<i>noké</i>	‘move’
e. —	<i>itfá</i>	<i>itfí</i>	‘plant’
f. <i>uku</i>	—	<i>uké</i>	‘rain’
g. <i>wiri</i>	<i>wirá</i>	<i>wiré</i>	‘stand’
h. <i>tfoʔi</i>	<i>tfoʔá</i>	<i>tfoʔi</i>	‘extinguish’
i. —	<i>osá</i>	<i>osé</i>	‘write’
j.—	<i>kimá</i>	<i>kimé</i>	‘cover with blanket’

38.3.3.2 *Change of State Predicates*

In addition to the valence stem allomorphy of unstressed roots, CR has a second valency-increasing stem allomorphy system, marked through thematic suffixes and stress shifts with a semantically defined set of predicates. The Guarijío cognates of these predicates are characterized as denoting an irreversible change of state or condition, generally produced by some kind of contact, with prototypical intransitive-causative pairs “break,” “twist,” “torn apart,” “shatter,” and “spill” (Miller 1996: 153). This class of verbs shares morphological traits across the UA language family: Heath (1978) reconstructs a morphological class of verbs for Proto-UA and the intermediate protolanguages, Proto-Northern-UA and Proto-Southern-UA, and gives descriptions of this class of verbs in Southern Paiute, Mono, Luiseño, Cupeño, Cahuilla, Serrano, Hopi, Tepiman, Rarámuri, and Nahuatl. This class is composed of disyllabic roots with thematic variation, and it includes verbs denoting events of physical change of state, even though the actual verbs may not necessarily be cognate from one language to another.

In CR, the class of change of state predicates is composed of unstressed roots that are marked as intransitive, transitive, or applicative through the presence or absence of transitive suffixes and specific stress patterns. All change-of-state verb roots are disyllabic, conforming to the pan-UA pattern. Intransitive stems have no transitivity suffixes and have fixed stress: stress is in the second syllable, whether a stress-shifting suffix (12a) or a stress-neutral suffix (12b) is attached.

(12) Intransitive change of state verbs

- a. tʃíhá-ma ré tʃíba
 scatter-FUT.SG DUB goats
 ‘The goats will scatter’
 ‘Se van a desparramar las chivas’
 [SFH 07 1:17/el]
- b. ma tʃíhá-ri tʃíba
 already scatter-PST goats
 ‘The goats already scattered’
 ‘Ya se desparramaron las chivas’
 [SFH 07 1:17/el]

Transitive stems, on the other hand, are marked through a transitive suffix, and are also sensitive to the stress properties of suffixes: stress is shifted to the transitive suffix, the third syllable, when combined with a stress-shifting suffix ((13a) and (13c)), but have final root (second syllable) stress when attaching a stress-neutral suffix ((13b) and (13d)).¹⁶

¹⁶ When bearing stress, the transitive suffix often causes the final root vowel to align in color with the first root vowel (e.g. *kuʔru-ná-ma* / *kuʔri-ná-ma* / ‘turn-TR-FUT.SG’).

(13) Transitive change of state verbs

- a. *tʃiha-ná-sa* *napátʃi*
 scatter-TR-CONDBlouse
 'If she scatters the blouses...'
 'Si desparrama las blusas'
 [SFH 07 1:17/el]
- b. *píri=mi* *orá* *tʃihá-ni-ri* *namúti*
 why=2SG.NOM why scatter.TR-TR-PST things
 'Why did you scatter the things?'
 'Por qué desparramas las cosas?'
 [SFH 07 1:17/el]
- c. *ma* *ripu-ná-ma* *ré* *kusí*
 now cut-TR-FUT.SG DUB wood
 'She'll cut the wood now'
 'Ya va a cortar la leña'
 [RF 04 verbs/el]
- d. *serútʃo* *ripú-ni-ri* *kusí*
 saw cut-TR-PST wood
 'The saw cut the wood'
 'El serrucho cortó la leña'
 [BFL 08 1:25/el]

Finally, applicative stems are marked through the addition of a transitive suffix and fixed final root (second syllable) stress. As shown in the next examples, the applicative stems of change of state verbs have fixed second syllable stress, whether they combine with a stress-shifting suffix (the future singular (14a)) or with a stress-neutral one (past (14b)).

(14) Applicative change of state verbs

- a. *nihé* *mi* *tʃiwá-ni-ma*
 1SG.NOM 2SG.ACC tear.APPL-TR-FUT.SG
 'I'm going to tear it for you'
 'Yo te lo voy a trozar'
 [SFH 07 1:21/el]
- b. *nihé* *ma=mi* *tʃihá-ni-ri* *napátʃi*
 1SG.NOM already=2SG.ACC scatter-TR-PST blouses
 'I scattered your blouses'
 'Ya te desparramé las blusas'
 [SFH 07 1:21/el]

The differences between the different kinds of stems are schematized in Table 38.2.

Table 38.2 Change of state predicates and thematic alternations

	<i>Stress-neutral</i>	<i>Stress-shifting</i>	<i>Properties of the stem</i>
<i>Intransitive</i>	<i>kasi-ri</i>	<i>kasi-ma</i>	Fixed second syllable stress No transitivizer suffix
<i>Transitive</i>	<i>kasi-ni-ri</i>	<i>kasi-ná-ma</i>	Transitivizer <i>-na</i> suffix Stress shifts
<i>Applicative</i>	<i>kasi-ni-ri</i>	<i>kasi-ni-ma</i>	Fixed 2nd syllable stress Transitivizer <i>-na</i> suffix

38.3.3.3 Causative and Applicative Suffixation

In addition to stem alternations and thematic suffixes, CR may also mark argument structure changing operations through multiple transitivizer, causative, and applicative suffixes, some of which have limited or no productivity. One of these suffixes is the transitivizer/pluractional *-tʃá*, an unproductive stress-shifting marker that combines with change of state predicates (the same predicates that take the *-na* suffix described earlier in the chapter). This suffix encompasses two senses: (i) an action is performed repeatedly (the succession or discernible, discrete events) and/or (ii) more than one entity is affected by an event.¹⁷ The contrast between the *-na* suffix and the pluractional transitive *-tʃá* suffix is exemplified in (14b, 14c). The stress-shifting transitivizing *-bu* suffix, which is also unproductive and lexically restricted, is exemplified in (15d, 15e).

(15) Restricted transitivizing suffixes

Basic (intransitive) construction

- a. *nihé* *kuʔrí-ma*
 1SG.NOM turn-FUT.SG
 ‘I will turn (on my own axis)’
 ‘Voy a dar vuelta (en mi propio eje)’
 [SFH 05 1:140/el]

Transitive construction

- b. *nihé* *kuʔru-ná-ma*
 1SG.NOM turn-TR-FUT.SG
 ‘I will turn it (on its own axis)’
 ‘Le voy a dar vuelta (en su propio eje)’
 [BFL 05 1:187/el]

¹⁷ Non-pluractional uses of the transitive *-tʃá* suffix can also be found, e.g. *rata-bá-tʃá-ma* ‘heat. ip-INCH-TR.PL-FUT.SG’ ‘will heat up the food’ [LEL 06 4:151/el].

Pluractional transitive construction

- c. *ma=ni* *kuʔrí-tʃí-ma*
 already=1SG. turn-TR.PL-FUT.SG.NOM
 'I will now turn it several times (on its own axis)'
 'Ya le voy a dar muchas vueltas (en su propio eje)'
 [BFL 05 1:187/el]

Basic (intransitive) construction

- d. *towí* *ma* *mó-ri*
 boy already go.up-PST
 'The boy already went up'
 'Ya se subió el niño'
 [BFL 06 4:189/el]

Transitive construction

- e. *ma=ni* *mo-bú-ru*
 already=1SG. go.up-TR-PST.PASS
 NOM
 'I was already taken up'
 'Ya me subieron'
 [BFL 06 4:189/el]

In addition to these restricted transitivity markers, CR has three applicative suffixes of limited distribution, *-ni*, *-si*, and *-wi*. These suffixes increase the valency of the verb, adding a benefactive argument, and are lexically conditioned by the roots to which they attach.¹⁸ The contrast between a basic, two-place base predicate and its applicative derivation with the *-ni* suffix is exemplified in (16).

(16) Restricted applicative suffix

Basic construction (two-place predicate)

- a. *ne* *ma* *wí-ma* *sunú*
 1SG.NOM now harvest-FUT.SG corn
 'I'll harvest corn now'
 'Ya voy a pizcar maíz'
 [LEL 06 4:151/el]

Applicative construction (three-place predicate)

- b. *wí-ni-mo=n* *orá* *ne* *jé-ra* *sunú*
 harvest-APPL-FUT.SG=1SG.NOM CER 1SG.NOM mother-POSS corn

¹⁸ It is possible to speculate that these suffixes at some point had additional semantic functions, since the cognates of these suffixes in River Guarijío encode more than one meaning (Miller 1996: 160–1). I have not, however, found any evidence that the CR applicatives introduce any semantic roles other than the benefactive/malefactive.

[BFL 06 5:146/el]

(17) CR causative construction

a. *ne* *mi* *rimé-ni-ra*
1SG.NOM 2SG.ACC make.tortillas-APPL-POT
'I can make you tortillas'
'Yo te hago tortillas'
[BFL o8 1:161/el]

b. *mi=ni* *ne* *ono-rá* *rimé-ni-ti-ma*
 2SG.ACC=1SG.NOM 1SG.NOM father-POSS make.tortillas-APPL-CAUS-FUT.SG
 ‘I will make you make tortillas for my dad’
 ‘Te voy a hacer que le hagas tortillas a mi papá’
 [BFL o8 1:161/el]

The productive applicative *-ki* suffix, on the other hand, is a stress-neutral suffix that introduces a benefactive or malefactive argument to one-place or two-place predicate. In (18b), the Applicative introduces a benefactive argument as an unmarked object (*mi* “2sg.acc”) to a basic transitive predicate (an argument which would be expressed through a postpositional phrase in a non-applicative construction).

Basic construction (two-place predicate)

- a. *ma=ni* *rata-bá-tʃɪ́-ki* *koʔwá-ami*
 already=1SG.NOM heat-INCH-TR.PL-PST.1 eat-PTCP
 ‘I already heated up the food’
 ‘Ya calenté la comida’
 [BFL o8 1:20/el]

Applicative construction (three-place predicate)

- b. *ne* *mi* *baʔwí* *rata-bá-tf̃-ki-raʔ*
 1SG.NOM 2SG.ACC water heat-INCH-TR-APPL-POT
 ‘Shall I heat the water for you?’
 ‘Te caliento el agua?’
 [BFL o8 1:21/el]

While instrumental prefixation and denominal derivation are processes that have a defined place within a single domain in the hierarchical, layered structure of CR verbal morphology, the Inner Stem, argument structure changing devices are distributed over the verbal complex, displaying their characteristic phonological and morphophonological properties. Most importantly, and as discussed next, the distribution of the multiple valence increasing markers over several verbal domains allows them to display certain type of interactions that are of typological and theoretical interest.

38.3.4 Derivation within the Agglutinating Verbal Complex

Three kinds of evidence converge in identifying zones or domains within the verbal complex: (i) morphotactics; (ii) the meanings of the categories expressed in each domain; and (iii) phonological phenomena that correlate with morphological boundaries. There are twelve suffix positions, grouped together with roots and stems into six concentric verbal zones or layers: the Inner Stem (the input to suffixation), the Derived Stem, the Syntactic Stem, the Aspectual Stem, the Finite Verb, and the Subordinate Verb. Each of these levels is semantically, morphotactically, and morphophonologically motivated. Crucially, derivation is concentrated on the innermost layers, the Inner Stem, the Derived Stem, and the Syntactic Stem.

The Derived Stem, the first identifiable layer of verbal suffixation, is a domain that contains semantically restricted, unproductive derivational suffixes, including the transitivizing *-na* and *-t f̃a* suffixes described in the previous section, which interact with stem allomorphy taking place at the Inner Stem level. Morphophonologically, this stem level is characterized by: (i) marking the non-concatenative imperative singular as final stress of this level; (ii) morphologically conditioned vowel lengthening (induced by the past passive suffix); and (iii) morphologically induced stress shifts.

A Derived Stem is the input to the next stage of the construction of the CR verb, the Syntactic Stem, which includes the valence-increasing suffixes described in Section 38.3.3.3. Morphophonologically, these suffixes form a coherent zone since they are stress-neutral and part of the domain of round harmony (that is, no later suffixes are targets for the harmony processes). An important property of these suffixes is that they may redundantly be added to stems already specified as transitive/causative or applicative in the Inner Stem or Derived Stem domains, a case of multiple exponence. Originally defined by Matthews (1974), multiple exponence involves the superfluous multiple marking of a single feature, bundle of features, or derivational category by multiple morphological layers (Caballero and Harris 2012). The examples in (19) show how the applicative is redundantly marked: there are two markers of applicative in (19b, 19d, 19f) but

only one benefactive or malefactive argument introduced in each case. These forms are equivalent to their counterparts with only one applicative exponent (19a, 19c, 19e).

(19) Optional applicative suffixation of applicative stems

- a. *ma=ne* *mi* *suwé-ri* *remé*
 already=₁SG.NOM ₂SG.ACC finish.APPL-PST tortillas
 'I already finished (ate) up your tortillas'
 'Ya me acabé tus tortillas'
 [SFH 05 1:119/el]
- b. *ma=ni* *mi* *suwé-ki-ri* *remé*
 already=₁SG.NOM ₂SG.ACC finish.APPL-APPL-PST tortillas
 'I already finished (ate) up your tortillas'
 'Ya me acabé tus tortillas'
 [LEL 06 5:123/el]
- c. *osí-ma*
 write.APPL-FUT.SG
 'She will write him (a letter)'
 'Le va a escribir (una carta)'
 [BFL 06 2:98/el]
- d. *osí-ki-ma*
 write.APPL-APPL-FUT.SG
 'She will write him (a letter)'
 'Le va a escribir (una carta)'
 [BFL 06 2:98/el]
- e. *roné-ma*
 boil.APPL-FUT.SG
 'He will boil it for her'
 'Se la va a hervir'
 [BFL 06 2:101/el]
- f. *roné-ki-ma*
 boil.APPL-APPL-FUT.SG
 'He will boil it for her'
 'Se la va a hervir'
 [BFL 06 2:101/el]

This semantically redundant interaction between different valence-changing morphological processes has been scarcely documented to date (Caballero and Harris 2012).¹⁹ CR multiple exponence of applicatives and causatives is at odds with one of the proposed diagnostic criteria of derivational vs. inflectional morphology, namely the property of

¹⁹ Within Uto-Aztecan, Classical Nahuatl also displays similar causative ME patterns (Launey (2002: 303), citing Carochi (1645), though Launey does not explicitly analyze patterns of causative doubling as semantically redundant; for discussion, see Caballero and Harris (2012)).

recursivity. Given widely held, but debated assumptions about the nature of expression of morphosyntactic features (namely, that inflectional information is marked only once in complex forms), it is expected that recursivity is possible for derivation, but not for inflection (Booij 2000). CR does possess instances of recursivity of argument changing morphology (e.g. where each causative marker adds an additional causer argument (Caballero 2008)), but these patterns coexist with the truly redundant ones, while none of the inflectional markers displays any similar redundancy.

Finally, despite having mostly a fixed position, there are no discontinuous dependencies across suffix positions in CR (and arguably elsewhere in UA), as attested in position class morphologies (Simpson and Withgott 1986, Stump 1992, Inkelas 1993, Good 2003). Instead, CR affix ordering conforms to principles of semantic compositionality (scope; Baker 1985b, Alsina 1999, Rice 2000), where affix order permutations largely correspond to changes in semantic interpretation (Caballero 2010). CR affix order is also consistent with the proposed cross-linguistic tendency of more semantically relevant markers to be closer to the root, and the proposed correlation of relevance with degree of phonological fusion, where (inner) derivational morphological exponents are expected to be more tightly fused phonologically with the root than (outer) inflectional affixes (Bybee 1985: 97).

38.4 CONCLUSION

In this chapter I have provided an overview of three major types of derivational processes in CR—instrumental prefixes, denominal verbalizing morphology, and a complex system of argument structure changing derivation marking—with a general characterization of their semantic functions and the variety of processes involved in their morphological expression. In this overview I have assumed that “[i]nflection differs from derivation in that derivation is a lexical matter in which choices are independent of the grammatical environment,” whether this environment is either syntactic or morphological (Bickel and Nichols 2007: 169). Other criteria proposed in the literature for distinguishing derivation from inflection characterize the CR and related UA constructions under discussion, including generality over the lexicon, relative order of morphemes (with derivation ordered before inflection, inner derivational markers being more tightly fused phonologically to the root than outer inflectional ones), and lack of semantic transparency.

Within the larger context of the CR verbal morphological system, the processes described here, together with other word-formation mechanisms that were left out of the scope of this chapter, fall into defined domains. These domains are motivated by independent semantic, morphotactic, and morphophonological evidence. Situating derivation mechanisms within this structure allows understanding certain properties and limits of their interaction, including the redundant marking of argument structure changing morphology.

CHAPTER 39

MATAGUAYAN

VERÓNICA NERCESIAN

39.1 THE MATAGUAYAN LANGUAGE FAMILY

MATAGUAYAN is a Southern Amerindian family comprising four languages: Chorote (or Manjui), Nivacle (or Chulupi or Ashluslay), Maka and Wichi (or Mataco) (see Swadesh 1959: 22, Najlis 1984: 5, Fabre 2012 [2005]: 1). Other known alternative names for the family are also found in the literature: “Matacoan” or “Mataco” (Loukotka 1968: 53, Greenberg 1987: 73),¹ “Mataco-Mataguayan” (Tovar 1951: 400, 1961, 1964), and “Mataco-Maka” (Kaufman 1990: 46).

Based principally on the lexical-statistical calculations made by Tovar (1964), Fabre (2012: 3) proposed a tentative internal classification of the Mataguayan family into two sub-branches: I. Wichi; II. Chorote, Nivacle, and Maka. He holds that Wichi seems to be closely related to Chorote from the lexical viewpoint; however, similarities on some morphosyntactic structures like deictic classifier systems induce him to group Chorote, Nivacle, and Maka together and Wichi separately. According to Swadesh (1959: 22), the internal divergence within this family could have taken place, at least, seventeen centuries ago. Nevertheless, as Fabre (2012) maintains, genetic ties among Mataguayan languages should be revised as more detailed documentation of these languages is obtained.

A genetic relationship between both Mataguayan and Guaykuruan has also been suggested, both constituting a larger group named Macro-Guayckuru or Mataco-Guaykuru (see Greenberg 1956, 1987, Viegas Barros 1993). Greenberg (1987) claims the Macro-Guaykuru group belongs to Macro-Panoan, within the large Ge-Pano-Carib,

¹ The term ‘Mataco’ was given by the non-indigenous population and many scientists to refer not only to the Wichi language and people, but also to the language family. However, native speakers reject this name because it has a pejorative sense, and they self-denominate using the native term Wichi, which refers to the language as well as the people.

coming together with Pano-Tacana and Moseten. However, this hypothesis has not yet been proved. Indeed, since the end of the 19th century, the first linguists that studied the Chaco languages (Lafone Quevedo 1895, 1896, 1912, Tovar 1951) developed their research mainly based on the strong conviction that the Chaco languages share features stemming from their contact, regardless of their different origins and their own particularities (Comrie et al. 2010: 85). Current typology-areal research on the Gran Chaco languages in South America have found prime evidence supporting the hypothesis that the languages spoken in the geographical region of the Gran Chaco constitute a linguistic area, which includes not only Mataguayan and Guaykuruan but also languages from other families (see Comrie et al. 2010, Golluscio and Vidal 2009/10).

39.1.1 Population and Location

The geographical extension of the Mataguayan languages is rather wide and spans three countries in the South American lowlands. They extend across northeastern Argentina, southeastern Bolivia, and southwestern Paraguay, and belong to the Chaco region. The Gran Chaco is a vast geographical area covering approximately 1,000,000 square kilometers. Mataguayan languages share the Gran Chaco with seven linguistic families (Tupi-Guarani, Guaykuru, Lule-Vilela, Arawak, Zamucoan, and Enlhet-Enenlhet or Maskoy) and the two isolated languages Besiro or Chiquitano and Guató (this last one considered extinct, as is the Guaykuruan language Abipon). The Chaco is a multilingual region, where forty native peoples live nowadays speaking, at least, twenty-nine languages (Golluscio and Vidal 2009/10: 5-ss).

The Mataguayan family shows a significant degree of vitality. Within the family, Wichi is the language with the broadest number of native speakers: 33,990 (INDEC 2004–2005) distributed in the Argentinean provinces of Formosa, Salta, and Chaco; and 4,115 (ORCAWETA 2011: 18) on the margins of the Pilcomayo River from Yacuiba (South) to Villamontes (North), in the county of Tarija, Bolivia. Wichi is followed by Nivacle, with 11,867 speakers (DGEEC 2003) spread across the Hayes and Boquerón counties of Paraguay, and approximately 553 in the town of Tartagal (province of Salta) and Las Lomitas (province of Formosa) in Argentina. Less numerous is the Chorote group, with 1,684 (INDEC 2004–5) speakers in the counties of Rivadavia (on the Pilcomayo River) and San Martín, province of Salta, Argentina, and about 1,313 speakers, according to Bareiro (2006, cited in Fabre 2012: 37), in the Boquerón county in Paraguay. Lastly, Maka has the smallest group of native speakers, with 1,267 people (DGEEC 2003) inhabiting the counties of Hayes and Qemkuket and north of Asunción city in Paraguay.

39.1.2 Linguistic Sources

Historical linguistics on the Chaco is still quite elementary compared with studies on other languages and families around the world, as Braunstein and Vidal (forthcoming)

hold. According to these authors, one of the reasons is that “some languages have yet to be documented thoroughly, in extensive dictionaries or comprehensive reference grammars, and most of what we currently know about others is incipient” (Braunstein and Vidal forthcoming). The Mataguayan family is no exception. There are only two studies specifically focused on the Mataguayan family: Najlis (1984) and Viegas Barros (2002), and circumscribed to the Proto-Mataguayan phonological inventory. A third paper makes a comparison between the laterals and labialized consonants among Mataguayan languages (Gerzenstein 2004). In sum, the grammatical study of this family is truly inchoate. Thus, this chapter is one of the first examining a specific grammatical issue, and it should be considered as such.

The sources consulted for this study consist of primary (the author’s own fieldwork) and secondary data (grammatical descriptions and vocabularies). On Chorote, a brief overview grammatical description (Gerzenstein 1978, 1983) and a 1,500-word vocabulary (Gerzenstein 1979, 1997a) were the main sources. For verbal derivation, particularly causativization, Carol (2009/2010) was consulted. On ethnobotanic studies, the work of Scarpa (2007, 2010), Arenas and Scarpa (2007) were most helpful. Some particular Chorote examples were scrutinized and offered (from their own fieldwork) by my colleagues Paola Pacor and Gustavo Scarpa. Regarding Maka, the grammatical description of Gerzenstein (1994), the study on negation (Gerzenstein 2002) and the papers on word formation in the ethnobotanic domain by Messineo and Tacconi (2010) and Messineo and Cúneo (2010) were the major sources. Occasionally, the bilingual (Maka-Spanish) vocabulary of Gerzenstein (1997b) was consulted too. As to Nivacle, a doctoral dissertation providing a grammatical description by Stell (1987) is the main documentation of the language so far. Alejandra Vidal, who is documenting and studying this language, assisted with some doubts on the Nivacle data. Finally, the Wichi data stem mainly from my own fieldwork (conducted since 2002 in the Formosa and Chaco provinces in Argentina, and, recently, in the Tarija county in Bolivia), and several studies (Nercesian 2007, 2008, 2009/10, 2011a, 2011b). Further key sources were Vidal and Nercesian (2005a, b, 2009a, b) and Nercesian and Vidal (forthcoming).

39.2 WORD-FORMATION PROCESSES

Mataguayan languages display a general uniformity with respect to word-formation processes. They all use derivation, composition, and incorporation² to form new lexical items. Wichi also uses onomatopoeia to create ideophones referring, basically, to birds

² As a word-formation process, incorporation (Type I, following Mithun (1984))—more studied in Wichi (Nercesian 2011a, b, 2012)—is a regular and systematic way to form a new verb, which even denotes a unitary concept. Only two verbs are used in this word formation process: *iwuye* ‘to do’ and *yenlhi* ‘to make’. Besides, only nouns can be incorporated, which may be alienable or inalienable, basic or derived, native or borrowed, and denote things, people, and, less frequently, body parts. The semantic role of the incorporated noun is not fully transparent. But, when it is, it may be a patient and,

and some intransitive activity verbs. In terms of productivity, derivation ranks first, followed by composition and then incorporation.

Inalienable nouns in this family require possessor prefixes to be attached to the base and to be recognized as a word by the speakers.³ In this sense, possession inflection proves a basic condition for word formation. However, inflection does not derive the base or create a new concept.

Some combinations of different word-formation processes are allowed in Mataguayan languages, although they are not highly common. Some examples from Wichi are shown in (1)–(6).⁴

- (1) Incorporation + derivation
iwu + 'elh-a = hi
₃SBJ:do+other-IC = LOC
 'He/she changes.'
- (2) Nominalization+incorporation
iwu + t'alh-yaj-a
₃SBJ:do + request-NMLZ-IC
 'He/she prays.'
- (3) Composition + incorporation
toyenkatmeklhoka
to-yen+ka-timek + lhok-a
₁SBJ.PL.IN-make + CL-hook+food-CI
 'we make bait'
- (4) Onomatopeia+reduplication
wowo [wo'wo] 'be dumb'
m'alh-m'alh ['maɬ'maɬ] 'greater wagtail-tyrant' (*Stigmatura budytoides*)
ch'ay-ch'ay [č'aj'č'aj] 'crested gallito' (*Rhinocrypta lanceolata*)
- (5) Onomatopeia+derivation: -*taj* 'Augmentative'
kalaktaj [qalaq'taχ] 'white-necked heron' (*Ardea cocoi*)
ch'estaj [č'es'taχ] 'barn owl' (*Tyto alba*)
hakhaktaj [hãq'haq'taχ] 'bare-faced ibis' (*Phimosus infuscatus*),
 'white-faced ibis' (*Plegadis chihi*)
- (6) Onomatopeia+reduplication+derivation
kow-kowtaj [qowqow'taχ] 'dark-billed cuckoo' (*Coccyzus melacoryphus*)
whu-whutaj [wũwũ'taχ] 'squirrel cuckoo' (*Piaya cayana*)

less frequently, an instrument. After the incorporation of the noun, transitive verbs become intransitive predicates.

³ Mataguayan languages are head-marking and tend to polysynthesis and agglutination.

⁴ I have followed the orthographic transcriptions (and phonological, when applicable) used in the sources by the authors. Some roots in Nivacle have been documented without the personal markers using “-” before the root, and I keep this notation. The legend “no reference” for the language’s forms means that there is no mention of this derivative in the sources consulted for each language.

39.3 DERIVATION

Different derivational processes are employed in the family: suffixation, prefixation, prefixation–suffixation, conversion and reduplication, though with different degrees of productivity and diffusion within the family. Full and (postposing) partial reduplication, for example, are both found in Nivacle as a verbal derivational process (i.e. *klopop* ‘it switches on’ > *klopopop* ‘it is luminous, bright’; *jakut* ‘(s)he is black’ > *jakutjakutxum* ‘(s)he whitens’; Stell 1987: 355–6), whereas in Wichi (Maka and probably also Chorote) it is mainly present in ideophonic names of birds (see example (4) above). Likewise, conversion (N→V), clearly attested in Wichi, is apparently missing in Nivacle, instead featuring several verbalizer suffixes (cf. 39.3.2.1). Prefixation–suffixation is another highly productive derivational process in the family. It creates, on the one hand, nouns expressing the lack of what is denoted by the base (i.e. Cho: *la’a=itsjemla-je* ‘without wife, single’ (Gerzenstein 1978: 81);⁵ Wi: *laj-la-wekw-a* ‘without owner, free, loose’ (Nercesian 2011a: 272)); and, on the other hand, verbs’ antonyms (i.e. Wi: *ni-is-a* ‘it is bad,’ lit. ‘it is not good’; *ni-lofwel-a* ‘it is slow,’ lit. ‘it is not fast’). Derivational prefixation forms reflexive predicates (i.e. Wi: *n’-lhi-lej* ‘I wash myself’ (Nercesian 2011a: 328)).

Nevertheless, suffixation is undoubtedly the most productive process for deriving words in the Mataguayan family. The fact that derivational processes are strongly suffixing is not due to a morphological constraint in the languages. Indeed, from a typological point of view and following inflectional morphology, Mataguayan languages tend to be equally prefixing and suffixing. Derivational morphemes are often easily recognizable owing to the low degree of fusion and a relatively high degree of agglutination in Mataguayan languages. In fact, more than one derivational suffix on a base is possible, rendering multiple suffixation derivation (cf. 39.3.1.8 and 39.3.2.3).

Nouns and verbs are the most likely to be derived, being also the word-classes usually created by derivation, involving transposition or not. Conversely, derived adverbials are conspicuously less common. In fact, special suffixes to derive adverbs were not found either in my own primary data from Wichi or in the secondary sources of the other languages. Still, some nominal and verbal categories (like demonstratives, tense, or directionals) combine with adverbs, nouns or conjunctions to create an adverb. For example, *ikana* ‘here’ < *ika=na* [location=DEM near.horiz]; *hotetsu* ‘like that (towards there)’ < *hote=tsu* [COMP=DEM.to_there]; *pajche* ‘a long time ago’ ‘in the old days’ < *paj=che* [then-DIR.in_extension]; *fwalanajhi* ‘yesterday’ < *fwala=najhi* [sun/day=DEM. near.past]; *toj’le* ‘there’ < *toj=l’e* [CONJ=DEM.far.horiz]. Hence, the analysis of these processes of adverbial formation as derivation may be rather weak. In Mataguayan languages, nouns, verbs, and adverbs are the major word classes; adjectives do not constitute

⁵ Abbreviations for the languages’ names: Cho = Chorote; Ma = Maka; Ni = Nivacle; Sp = Spanish; Wi = Wichi.

a clear-cut single word-class. Rather, they form a specific group of intransitive verbs that denote properties and states. Nouns and verbs differ from adverbs basically in that the latter do not inflect; the distinction between nouns and verbs is more flexible (see Vidal and Nercesian 2005a): both share tense suffixes, evaluative morphology and negatives without triggering a change of word-class. Through conversion, indeed, denominal verbs may carry verbal pronominal categories. However, possessive prefixes, gender (only in Maka) and number are categories not shared by these two parts of speech.

This study concentrates on the derivational processes through suffixing in nouns and verbs, this being the most productive in the family. Certainly, it does not cover all derivational processes in the group, but the key ones as to productivity and diffusion.

39.3.1 Deriving Nouns

Nominal derivation is effected exclusively through affixation, suffixation being by far the main process to form nouns. The most common semantic categories in nominal derivation are “agent,” “patient,” “action,” “instrument,” “locative,” “evaluative,” and “vegetable class.” Transposition of word-class can result from processes forming “agent,” “patient,” or “action” nouns; these are deverbal nominalizations. “Locative,” “evaluative,” and “vegetable class” derivations do not involve a word-class change. Some derivational processes may trigger a change in the nominal class (alienable/inalienable) of the base. Thus, agent, instrument, and vegetable class derivations, on the one hand, create alienable nouns; and patient, action, and instrument (only Maka) derivations, on the other hand, create inalienable nouns. The nouns derived through locative, augmentative, and diminutive are alienable or inalienable depending on the class to which the base belongs. However, derived locative nouns tend to be used in the undefined possessor form. The alienable/inalienable distinction is semantically based, so the change of the nominal class is not irrelevant for the system. Alienable nouns refer to objects, the physical world, vegetation, and animals; inalienable nouns, on the other hand, refer to parts of the body, kinship, and (generally personal) body-related objects.

Polysemy of the suffixes is quite uncommon in Mataguayan languages. Only agent (cf. 39.3.1.1) and instrumental (cf. 39.3.1.4) derivatives have more than one meaning, and there is some interesting semantic overlapping between “agent” and “instrument,” especially when the referent is animate.

Each nominal derivative selects a different plural suffix. Mataguayan languages have a set of suffixes encoding plural number (the singular being unmarked) that combine with different bases, thus forming subsets of nouns. The plural suffix that the derived noun selects is not necessarily the same as the one that combines with the non-derived base: *afwenche-y* ‘bird-PL,’ *afwenche-fwa(j)-s* ‘bird-DIM-PL’; *tente-y* ‘stone-PL,’ *tente-ta(j)-s* ‘stone-AUG-PL’; *hal’o-y* ‘tree-PL,’ *hal’o-wu-s* ‘tree-AG-PL,’ *hal’o-cha-lh* ‘tree-INS-PL’; *mati-lis* ‘mate-PL,’ *mati-kut[s]-es* ‘mate-AG-PL’.

Lastly, the same base (even Spanish loanwords) can be used with a number of different suffixes: *mati* ‘mate’; *mati-cha* [mate-INS] ‘mate straw’; *mati-wu* [mate-AG] ‘mate pourer’; *mati-kut* [mate-AG] ‘mate-drinker’; *fwaʔay* ‘carob fruits’; *fwaʔay-cha* ‘tool used to gather carob fruits’; *fwaʔay-wu* ‘carob weevil’; *fwaʔay-kut* ‘carob-eater.’

39.3.1.1. Agent Nouns

Derived agent nouns can be denominal or deverbal. In the latter case, derivation implies nominalization. The nominalizers that create agent nouns are formally alike among the four languages [-ta, -tsax, -nax]. The alternation between [-tsax] and [-nax] seems to be phonotactic, yet the constraint is not totally clear. Examples of deverbal nominalization are given in examples (7)–(10).

- (7) Chorote: -(V)*ta*
- | | |
|--|--|
| a. <i>amtije-ta</i>
to_talk-NMLZ
‘chatter-box’ | b. <i>wo(ʔo)-ta</i>
to_fish-NMLZ
‘fisherman’
(Gerzenstein 1978: 84) |
|--|--|
- (8) Maka: -(V)*nax* ~ -(V)*tsax*
- | | |
|--|---|
| a. <i>yixalkʼis-i-naX</i>
to_hunt-ep-NMLZ
‘hunter’ | b. <i>yi Yin-he-tsaX</i>
to_request-ep-NMLZ
‘pest’
(Gerzenstein 1994: 222) |
|--|---|
- (9) Nivacle: -(*xa*)*nax* ~ -*ax*
- | | |
|---|--|
| a. <i>ʔp-xanax</i>
to_cry-NMLZ
‘cry baby’ | b. <i>ʔlawiʃxan-nax</i>
to_sing-NMLZ
‘singer’
(Stell 1987: 215) |
|---|--|
- (10) Wichi: -*naj* ~ -*tsaj* ~ -*aj*
- | | |
|--|---|
| a. <i>(ta)yhi(n)-naj</i>
to_make_crafts-NMLZ
‘artisan’ | b. <i>(yi)ku(y)[n]-tsaj</i>
to_play-NMLZ
‘player’
(Nercesian 2011a: 270) |
|--|---|

The derived agent nouns refer to a person who often does or practices an activity, or usually displays some particular behavior. It belongs to the alienable class of nouns, hence not carrying the pronominal possessor prefix.

In addition to nominalizers, Chorote and Wichi have another suffix, -*wo* ~ -*wu*,⁶ to create denominal agent nouns referring to people who make or work with something, thus having specific skills. In Wichi, the suffix -*wu* is polysemic and conveys two

⁶ The suffix -*wu* could have developed from the verbal root *iwuʔe* ‘to do,’ which synchronically coexists as an independent verb in Wichi.

meanings: (i) somebody/thing that makes something (12a, 12b); and (ii) somebody/thing that lives, develops or reproduces somewhere (12c, 12d).

(11) Chorote: *-wo*

a. *impeniki-wo*

food-AG

'cook'

(Gerzenstein 1978: 84)

b. *inkakie-wo*

medicine-AG

'doctor'

(Pacor and Scarpa p.c.)

(12) Wichi: *-wu*

a. *pini-wu*

honey-AG₁

'bee'

c. *inot-wu*

water-AG₁

'water animal'

(Nercesian 2011a: 268–9)

b. *sapatu-wu*

shoe-AG₁

'shoemaker'

d. *hunatsi-wu*

night-AG₁

'nocturnal insect'

The same suffix is found in Maka and Nivacle but creating instrumental nouns rather than agent nouns (cf. 39.3.1.4). Wichi and Chorote use a different suffix for this. This semantic mismatch of the suffix *-wu*~*-wo* among the languages of the family (as an agent derivative in Wichi and Chorote and as an instrument in Maka and Nivacle) is noteworthy and not rare at all, since the semantic relation between “agent” and “instrument” may be close in some sense. An agent could serve as an instrument or as a means for a special achievement. According to Luschützky and Rainer (2011: 326), among the world's languages expressing both “agent” and “instrument” by means of word formation, the percentage of polysemous markers is considerably significant (48%). They argued that the reasons for the semantic change giving rise to polysemy may be a range of different pathways of change from a diachronic perspective (Luschützky and Rainer 2011: 327). In addition, as Booij and Lieber (2004) hold, “paradigmatic pressure” may be a motivation for the polysemy of the derivational markers. Diachronic research on agent and instrument derivational forms in Mataguan languages is necessary to elaborate hypotheses on their origin and the pathways of semantic change they might have followed.

A third agent derivational suffix in Wichi is *-kut*, which combines with nouns and refers to a person who feeds on something or someone who always eats X:

(13) Wichi: *-kut*

a. *w'ahat-kut*

fish-AG₂

'fish-eater'

(Nercesian 2011a: 269)

b. *pini-kut*

honey-AG₂

'honey-feeding'

Summing up, deverbal nominalizations create agent nouns referring to a person who usually does something, and denominal derivation, on the other hand, creates agent nouns referring to a person who makes or builds what the base denotes.

39.3.1.2 Patient Nouns

Derived patient nouns come from a verbal nominalization process that adds the suffix *-Vk ~ -Vč (the palatalized form in Nivacle, cf. (14a, 14b)).⁷ The derived noun refers to the result, effect or product of the event denoted by the base. Thus, the derived noun from *kill* means *dead person/animal*; and the one from *make* means *work*. See examples (16)–(19).

- (14) Maka: -ek ~ -ik
- | | |
|--|--|
| <p>a. <i>witi-su'un-ek</i>
INDEF.POSS-to_love-NMLZ
'person loved by someone'</p> | <p>b. <i>wit-ophet-ik</i>
INDEF.POSS-to_tie-NMLZ
'thing tied by someone'
(Gerzenstein 1994: 225)</p> |
|--|--|
- (15) Nivacle: -ač ~ -eč
- | | |
|---|---|
| <p>a. <i>-en-ač</i>
-to_love-NMLZ
'person loved by someone'</p> | <p>b. <i>-kakla^{xł}-eč</i>
-to_fry-NMLZ
'thing fried by someone'
(Stell 1987: 215)</p> |
|---|---|
- (16) Wichi: -ek
- | | |
|--|--|
| <p>a. <i>la-chefwen-ek</i>
3POSS-to_teach-NMLZ
'his/her student'
'my work'</p> | <p>b. <i>n'lhenek</i>
n'-yenlhi-ek
1POSS-to_make-NMLZ
(Nercesian 2011a: 279)</p> |
|--|--|

The suffix -Vk ~ -Vč is employed for the singular form. To express the plural of the derived noun this suffix is replaced by -hVy, which happens to be the same pluralization pattern as in the set of non-derived nouns that end in -ek (e.g. *lutek* 'arrow', *luthay* 'arrows'; *lhetek* 'head', *lhehey* 'heads').

(17) Maka (Gerzenstein 1994: 225) <i>witi-fen-hay</i> INDEF.POSS-to_aid-NMLZ:PL 'aided by somebody'	(18) Wichi (Nercesian 2011a: 279) <i>la-chefwen-hay</i> 3POSS-to_teach-NMLZ:PL 'his/her students'

Unlike deverbal agent nouns that are alienable (cf. 39.3.1.1), deverbal patient nouns belong to the inalienable class, hence requiring possessor markers, as examples

⁷ No reference for Chorote.

(14)–(18) show. The possessor prefix attached to the derived nominal base refers to the agent of the event denoted by the verb stem. The participant is the owner because he/she is responsible for the result, effect, or product. Syntactically, the deverbal noun can function as an argument but also as an attributive nominal predicate, despite carrying nominal inflection, like possessor, number, and gender (when applicable, as in Maka). These properties lead linguists to analyze it as a participial form or masdar (Haspelmath 2002: 230). Finally, I have observed that in Wichi patient nominalization applies exclusively to agentive verbs; this appears to be a semantic restriction to derive non-agentive verbs denoting states or non-causative change of states by this type of nominalization.

39.3.1.3 Action Nouns

A third nominalization process creates abstract nouns referring to actions. This is a highly productive derivational process that applies to intransitive, transitive, agentive, and some non-agentive verbal bases.⁸ (In a similar way to the patient nominalizer, the derivational suffix $*-yV\chi$ palatalizes in Nivacle as $-a\check{s}$.)

(19) Maka: $-(y)aX \sim -(y)eX$

a. *wit-ixiwey-aX*

INDEF.POSS-be_afraid-NMLZ

‘someone’s fear’

(Gerzenstein 1994: 219–220)

b. *wit-kétey-eX*

INDEF.POSS-to_discuss-NMLZ

‘someone’s discussion’

(20) Nivacle: $-(xa)ja\check{s} \sim -a\check{s}$

a. *-tanšan-xajaš*

-to_wish-NMLZ

‘wish’

(Stell 1987: 216)

b. *-taiji-jaš*

-be_good-NMLZ

‘goodness’

(21) Wichi: $-yaj$

a. *n’-chefwen-yaj-ay*

1POSS-to_teach-NMLZ-PL

‘my teachings’

(Nercesian 2011a: 276)

b. *a-t’ische-yaj*

2POSS-to_laugh-NMLZ

‘your laughing’

c. *asnam-yaj*

3POSS:be_blind-NMLZ

‘his/her blindness’

Regarding the data I have registered in Wichi, the nominalization of non-agentive verbs is not very productive. Some derivational processes by $-yaj$ are possible for the speakers but not very common—as *la-yotaj-yaj* [3POSS-be_fat-NMLZ] ‘fatness’—yet in other cases the derivation is completely unacceptable.

As occurs with patient nominalization, derived action nouns belong to the inalienable class and have possessor markers. But, unlike deverbal patient nouns, deverbal

⁸ No reference for Chorote.

action nouns behave morphosyntactically as a plain nominal base. They always function as arguments and inflect as any other noun.

39.3.1.4 Instrumental Nouns

Instrumental derivational suffixes, unlike the others explored so far, differ across the languages in the family.

As shown in Table 39.1, Maka and Nivacle share the same instrumental suffix *-wo*, which combines with nominal bases. A similar form, *-wo* (Cho) and *-wu* (Wi), functions as an agent derivative in Chorote and Wichi (cf. 39.3.1.1). The polysemy of *-wu* in Wichi does not include the category of “instrument.” We do not know whether the original meaning of the proto-form **-wV* was “agent,” “instrument”—or both—and, therefore, what type of semantic shift could have happened from the diachronic viewpoint. The semantic closeness between “agency” and “instrument,” especially with an animate referent, as explained at the start of this section, clearly allows this crossing of the suffix meaning in the family. In addition, Scarpa (p.c.) registered the suffix *-hiwo* as an instrument in *ikenisye-hiwo* ‘religious celebrations chanter,’ which is semantically and formally akin to the agent derivative *-wo*.

The derivation through the suffix *-cha* in Wichi creates nouns referring to an instrument, as in (22a, 22b), where the referent is inanimate; but also to an agent person when the referent is animate, as in (23a, 23b).

- (22) Wichi
- | | |
|---|--|
| a. <i>fwiy'et-cha</i>
cold/winter-INS
'winter coat' | b. <i>chemet-cha</i>
work-INS
'working tool' (i.e. machine, screwdriver) |
|---|--|
- (23) Wichi
- | | |
|--|---|
| a. <i>inot-cha</i>
water-AG.INS
'person responsible for getting water' | b. <i>w'ahat-cha</i>
fish-AG.INS
'person responsible for catching fish' |
|--|---|

(Nercesian 2011a: 271)

Table 39.1 Derivational instrumental suffixes

Language	Suffix	Word class of the base
Cho	<i>-hiwo</i>	N→N
Ma	<i>-wo'</i>	N→N
	<i>-laX</i>	V→N
Ni	<i>-(xa)wo</i>	N→N
	<i>-xat</i>	V→N
Wi	<i>-cha</i>	N→N

Indeed, derived words such as (23b) and similar ones like *pini-cha* [honey-INS] may be polysemic and refer to both the instrument (i.e. the tool used to catch fish or collect honey) and the agent (i.e. the person responsible for catching fish or collecting honey).

Some nominal bases can be derived in their plural form, as example (24a) from Wichi illustrates. In these cases the pluralized nominal base refers semantically to a general class of entities; thus it is conceptualized as singular and, consequently, can in turn be pluralized, (24b). Note that the plural suffix attached to the base and the derived noun diverge.

- (24) Wichi
- | | |
|--|--|
| a. <i>mawuwu-s-cha</i>
flea-PL-INS
'flea-killer' | b. <i>la-ka-mawuwu-s-cha-lh</i>
3POSS-CL-flea-PL-INS-PL
'his/her flea-killers' |
|--|--|
- (Nercesian 2011a: 271)

We do not have robust evidence about the origin of *-cha*. However, from a synchronic perspective, this bound form may combine with the possessive classifier *ka-*, which usually accompanies alienable nouns in possessive constructions, thus constituting a single word (i.e. *kacha* 'remedy'); or it may even combine with the third person possessor prefix *la-* as happens with inalienable nouns (i.e. *lacha* 'tool'). This inconsistency—behaving as an alienable and an inalienable stem—and the notable productivity of the bound form *-cha* make us analyze it as a derivative suffix possibly resulting from the grammaticalization of an old free form.

The suffix *-xat* in Nivacle is curiously similar to the instrumental nominalizer suffix in Guaykuruan languages, *-gat* in Pilaga (Vidal 2001: 105), *-Gat* in Mocovi (Gualdieri 1998: 157), *-aGat* in Qom or Toba (Messineo 2000: 148). Since the suffix *-xat* of Nivacle is absent in the other Mataguayan languages, we do not discard the possibility of this form being borrowed from a Gaykuruan language with which Nivacle could have been in contact. As in the Guaykuruan languages, the Nivacle derivative *-xat* is a nominalizer:

- (25) Nivacle
wamka-xat
to_wash-NMLZ.INS
'washing tool'
(Stell 1987: 212)

Apart from *-xat*, Nivacle has a denominal instrumental derivative, the suffix *-(xa)wo*, as shown in Table 39.1, whose origin remains unknown.

- (26) Nivacle
kača-wo
throwing_element-INS
'throwing bird-catching wire'
(Stell 1987: 213)

According to Stell (1987), the suffix *-wo* also combines with verbal bases, as seen in (27).

- (27) Nivacle
 kum-**xawo**
 to_work-INS
 ‘tool’
 (Stell 1987: 213)

A similar case shows that Maka has two different instrumental suffixes, one of which is a denominal derivative: *-wo*, (28); and the other a nominalizer: *-laX*, (29).

- (28) Maka
- | | |
|---|--|
| <p>a. <i>wit-inxeyi-wo</i>
 INDEF.POSS-nasal_mucus-INS
 ‘somebody’s handkerchief’</p> | <p>b. <i>yi-to(‘)-wo-ts</i>
 1POSS-eye-INS-PL
 ‘my glasses’
 (Gerzenstein 1994: 221)</p> |
|---|--|
- (29) Maka
- | | |
|---|---|
| <p>a. <i>wit-wet-i-laX</i>
 INDEF.POSS-to_climb-ep-INS
 ‘somebody’s ladder/staircase’</p> | <p>b. <i>wit-ots-hi-laX</i>
 INDEF.POSS-to_tread-ep-INS
 ‘somebody’s carpet/rug’
 (Gerzenstein 1994: 222)</p> |
|---|---|

As Gerzenstein (1994: 222) has argued, *-laX* combines with intransitive verbs. The semantics of the suffix is a general idea of instrument and finality: “an object useful to do something.” The source of *-laX* is unknown and it has not been found in either Mataguyan or a neighboring language so far.

39.3.1.5 Locative Nouns

There are two common and highly productive bound forms that create locative nouns in Mataguyan languages. One of them, *-hi* or *-xi*, refers to a place that serves as a container; the other one, *-w’et* ~ *-wet*, refers to the place where an activity takes place or the place for something. The locative *-xi* that attaches to the noun in Chorote indicates the location of the event, but does not derive the noun (Gerzenstein 1978: 82). However, Scarpa (2007) has documented a derived form that conflates both, *-hi* and *-wet*, where *-hi* could function as the same derivative in Wichi: *inot* ‘water’ > *inot-hi* ‘puddle, lake, canal’ (Wi); and *anat* ‘water’ > *anat-hi-wet* ‘spring’ ‘the place where the water is’ (Cho). In Maka, the bound form *-wet* combines with verbs, whereas in Nivacle and Wichi it combines with both nouns and verbs.

Language	-hi ~ -xi ~ -ši “the container of”	-wVt “the place of”
(30) Chorote	no reference	<i>anat-hi-wet</i> water-LOC?-LOC “spring” (Scarpa 2007)
(31) Maka	a. <i>wit-finka-xii</i> INDE.POSS-tobacco-LOC “somebody’s pipe” (Gerzenstein 1994: 221)	b. <i>wit-ithayi-wet</i> INDE.POSS-to_work-LOC “the place where people work” (Gerzenstein 1994: 221)
(32) Nivacle	a. <i>-u^xtu-xi</i> -urine-LOC “bladder” (Stell 1987: 222)	b. <i>-samku-wat</i> -excrement-LOC “latrine” (Stell 1994: 220) c. <i>-man^xta-wat</i> -to_live-LOC “the place to live” (Stell 1994: 220)
(33) Wichi	a. <i>to-kos-hi</i> INDE.POSS-plant-LOC “plant pot” (Nercesian 2011a: 288)	b. <i>to-wej-w’et</i> INDE.POSS-buttocks-LOC “chair” (Nercesian 2011a: 288) c. <i>to-lhataj-w’et</i> INDE.POSS-to_cook-LOC “kitchen” (Nercesian 2011a: 288)

As I have observed in Wichi, both locative forms have evolved as bound morphs from words frequently used in compounds. Synchronically, they are midway through composition and derivation. Many suffixes and prefixes, as noted in Olsen (2000: 901–2), come diachronically from the first or second member of a compound. A free form can serve as the base for a compounding pattern, and when this pattern has consolidated and becomes productive, the original constituent starts to depart from its correspondent free form or its meaning, developing as an affix. That is the case of the locative bound forms of Wichi, which seem to be in an intermediate stage of grammaticalization. On the one hand, the forms *-hi* and *-w’et* exist as free words: *ihi* ‘there is/are, to exist’ and *la-w’et* ‘3POS-house, home’ respectively. Morphophonological processes, which occur when affixes and clitics add to the base within the phonological word, are blocked on the morphological edge between the locative form and the base, and combine with nominal and verbal word-classes (generally, derivational affixes select one word-class). On the other hand, however, not only are the bound forms *-hi* and *-w’et* highly productive (also frequently used for innovations and neologisms), but they

also lost phonological material, have a recognizable recurrent meaning and the meaning of the new word is considerably predictable.⁹

39.3.1.6 Evaluative Morphology

Languages have the possibility of interpreting diminution and augmentation in affecting terms, in addition to purely objective terms, as noted in Wierzbicka (1980: 53), Szymanek 1988: 106), Stump (1993: 1), and Kortvelyessy (chapter 24, this volume), among others. Thus, in Mataguayan languages words derived through evaluative morphology may generally express an assessment of the size of the base noun's referent, as in (34a–34c) and (35a–35c), or some positive or negative feeling of the speaker towards an entity, as shown in (36) and (37) from Wichi.

- (34) Wichi
 a. *afwenche* b. *afwenche-fwaj* c. *fwenche-taj*
 ‘bird’ ‘little bird’ ‘big bird’
- (35) a. *tente* b. *tente-fwaj* c. *tente-taj*
 ‘stone’ ‘little stone’ ‘big stone’
- (36) *n'-lhose-fwaj*
 ₁POSS-daughter-DIM
 ‘my little daughter’ (with positive connotation)
- (37) *hin'u-fwaj*
 man-DIM
 ‘little man’ (with negative connotation)
 (Nercesian 2011a: 266)

The augmentative suffixes are formally similar in the family (see Table 39.2), which means that they come from a single proto-form. In contrast, the diminutive suffix differs among Chorote, Nivacle, and Wichi.

Table 39.2 Evaluative morphology suffixes

Language	Diminutive	Augmentative
<i>Chorote</i>	<i>-les/-lis</i>	<i>-tje</i>
<i>Maka</i>	no reference	<i>-taX</i>
<i>Nivacle</i>	<i>-xʔas</i>	<i>-tax</i>
<i>Wichi</i>	<i>-fwaj</i>	<i>-taj</i>

⁹ A similar grammaticalization process of the verb *ihi* ‘there is/are; exist’ into a locative/applicative suffix seems to be unfolding as a consequence of its frequent usage in serial verb constructions (Nercesian 2009/10).

The Nivacle suffix *-xɬas* could have developed from the inalienable word *xɬaos* ‘3POSS:son.’ If we consider what happened in Wichi, we may suspect that in Nivacle the diminutive suffix is the result of the grammaticalization of this inalienable word as part of a compounding pattern. In Wichi, besides the suffix *-fwaj*, there is another way of expressing diminution that consists of a compounding process with the inalienable root *-les* ‘offspring (son or daughter).’ The compound can refer to a baby animal, (38a), or a little thing, (38b). The diminutive meaning of the compound may be an innovation, as in (38b), where the Spanish loanword *platu* (Wi) < plato (Sp) was employed as a member of the compound that refers to a culturally new dish. In Nivacle, as well as in Chorote, the diminutive suffix derives the nominal base expressing the assessment of the size of the base noun’s referent; see (39)–(40).

(38) Wichi

a. <i>hu’u+les</i> hen +baby ‘chick’	b. <i>platu+les</i> plate+baby ‘little plate’

(Nercesian 2011a: 289)

(39) Nivacle

a. <i>watcačikla-xɬas</i> earring-DIM ‘little earring’	b. <i>tannuk-xɬas</i> cat-DIM ‘kitty’

(Stell 1987: 224)

(40) Chorote

- ipiata-lis*
corn-DIM
‘little corn’

(Scarpa 2007)

The productivity of the Wichi compound with the root *-les* is a favorable context for a possible grammaticalization, which could turn compounding into a derivational word-formation process. It is also interesting to compare the strategy of Wichi and Nivacle when using the word “son” or “offspring” to express diminution. The semantic connection between the meaning of “son” and “affection” is easy to establish.

In addition, the augmentative can create new words from an existent word or root that does not exist as a single word, as shown in (41)–(42) and (43), respectively, for Wichi.

(41) Wichi

a. <i>alhe</i> ‘iguana’	b. <i>alhe-taj</i> ‘cayman’

(42) Wichi

a. <i>to-ch’ute</i> ‘somebody’s ear’	b. <i>ch’ute-taj</i> ‘cup’

- (43) Wichi
 a. *kalaktaj* 'white-necked heron'
 b. *tsilik-kliktaj* 'American kestrel'
 c. *ch'estaj* 'barn owl'
 (Nercesian 2011a: 268)

Apart from the possibility of interpreting diminution and augmentation in affective and purely objective terms in Wichi, evaluative suffixes can express manner when combined with verbs, as in (44a, 44b) and (45).

- (44) Wichi
 a. *asinoj yik-fwaj* dog ₃SBJ:go-MAN 'The dog went slowly'
 b. *lhetek ichot-fwaj* ₃POSS:head ₃SBJ:be_red-MAN '...and the head [of *Takfwaj*] was reddish...'
 (Nercesian 2011a: 404–5)

- (45) Wichi
lawut wefw-taj wit pitaj-taj
 beehive_hole ₃SBJ:be_big-AUG CONJ ₃SBJ:be_long-AUG
 'they are very large and long beehive holes.'
 (Nercesian 2011a: 406)

Within the family, the use of the diminutive with verbal predicates is also present in Nivacle, but with a different suffix, *-kle* (i.e. *-klap'af-kle* 'be baldish'; *-jicxu-kle* 'be a tough guy'; Stell 1987: 339).

39.3.1.7 Vegetable Class

The extremely stimulating research on ethnobotany in the Gran Chaco by Arenas (2003), Arenas and Scarpa (2007) and Scarpa (2007, 2010) allows the examination of word-formation processes in this culturally central domain of flora. Mataguyan languages have a special set of suffixes that derive the base as they also classify the flora elements. Names of fruits generally end in *-V~ -V'* or *-Vx*; most of the names of wooden plants (trees, bushes, lianas, or creepers) are formed by deriving the base referring to the fruit with the suffix *-Vk* or *-Vkw*; and a community of wooden plants, using the suffix *-Vt* or *-chVt*.

Language	Fruits (FR)	Wooden plants (WD)	Community of wooden plants (COM)
(46) Chorote	a. <i>jwa'a</i> carob 'carob fruit (<i>Prosopis alba</i>)' (Scarpa 2007: 88)	b. <i>jwa'a-yuk</i> carob-CL _{wd} 'carob plant' (Scarpa 2007: 88)	c. <i>jwa'a-kyet</i> carob-CL _{com} 'community of carobs' (Scarpa p.c.)

(47) Maka	a	<i>lekit-i</i> pumpkin-CL _{fr} 'pumpkin fruit' (<i>Cucurbita maxima</i>) (Messineo and Tacconi 2010: 100)	b.	<i>lekit-uk</i> pumpkin-CL _{wd} 'pumpkin plant' (<i>Cucurbita maxima</i>)' (Messineo and Tacconi 2010: 100)	c.	<i>tsepeq-et</i> Devil's_grape-CL _{com} 'Devil's grape vineyard (<i>Cissus palmata</i>)' (Messineo and Tacconi 2010: 98)
(48) Nivacle	a	<i>axaj-ex</i> mistol-CL _{fr} 'mistol fruit' (<i>Ziziphus mistol</i>)' (Stell 1987: 209)	c.	<i>šekla-juk</i> carob-CL _{wd} 'carob plant (<i>Prosopis alba</i>)' (Stell 1987: 210)	d.	<i>šekla-čat</i> carob-CL _{com} 'community of carobs' (Stell 1987: 210)
(49) Wichi	a	<i>fwa'a</i> Prosopis_alba 'carob fruit' (<i>Prosopis alba</i>)' (Nercesian 2011a: 262)	b.	<i>fwa'a-yekw</i> carob-CL _{wd} 'carob plant' (Nercesian 2011a: 262)	c.	<i>fwa'a-chat</i> carob-CL _{com} 'community of carobs' (Nercesian 2011a: 262)

Besides the collective of wooden plants that refers to a community, the nouns that refer to fruits and plants can pluralize. In the case of wooden plants it does so through suppletion; see examples from Wichi where the suffix *-(y)ekw* is replaced by *-chey* (a special plural form used exclusively with this class of nouns).

(50) Wichi

a.	<i>fwa'a-yekw</i> carob-CL _{wd} 'carob plant (<i>Prosopis alba</i>)'	b.	<i>fwa'a-chey</i> carob-CL _{wd} .PL 'carob plants'

Other types of plants are named by compounding processes. Both derivation and composition are equally productive in this specific semantic domain of the lexicon.

39.3.1.8 Multiple Derivational Suffixation

A derived word may receive another derivational suffix, thus causing multiple suffixation. As I have found in Wichi, the agent derivative can take an action or patient nominalized base and derived locative nouns. See (51)–(53) respectively.

- (51) *n'-chefwen-yaj-wu-s*
1POSS-to_teach-NMLZ-AG-PL
'my teachers'
(Nercesian 2011a: 277)

- (52) *to-lon-ek-wu*
 INDEF.POSS-to_kill-NMLZ-AG
 ‘what announces death (said of a bird)’
 (Nercesian 2011a: 281)
- (53) *to-kos-hi-wu*
 INDEF.POSS-plant-LOC-AG
 ‘farm keeper’
 (Nercesian 2011a: 269)

The diminutive and augmentative suffixes can take as a base a noun derived by the action nominalizer, (54) and (55), and through the patient nominalizer, (56) and (57).

(54) <i>to-chefwen-yaj-fwaj</i> INDEF.POSS-to_teach-NMLZ-<u>DIM</u> ‘a little teaching’ (Nercesian 2011a: 280)	(55) <i>to-tichunha-yaj-taj</i> INDEF.POSS-to_think-NMLZ-<u>AUG</u> ‘big thinking’ (Nercesian 2011a: 277)
(56) <i>a-lon-ek-fwaj</i> 2POSS-to_kill-NMLZ-DIM ‘your little killing’ (it is little because of the size of the killed) (Nercesian 2011a: 281)	(57) <i>n'-lhen-ek-taj</i> 1POSS-to_make-NMLZ-<u>AUG</u> ‘my big work’ (Nercesian 2011a: 280)

As the examples show, more than one derivational suffix can attach to the same root in Wichi. We do not have sufficient data in the other languages of the family. Therefore, we are not able to state whether Nivacle, Maka, and Chorote have the same possibility as Wichi. However, it is likely that similar cases are to be found in these languages.

39.3.2 Deriving Verbs

Verbal derivation is mainly effected through affixation, with suffixation being the most productive and prefixation and prefixation–suffixation less so. On the one hand, prefixation is found in Maka, Nivacle, and Wichi in the derivation of reflexive verbs, which in turn decreases their valency. On the other hand, through prefixation–suffixation the negative is used in Mataguyan languages to form antonyms. Besides affixation, conversion is another process used to derive denominal verbs in Mataguyan languages. Partial and full reduplication are also found in Nivacle to form inchoative predicates from stative verbs. Comparing nominal and verbal derivational processes, a wider variety of ways to create nouns rather than verbs was attested.

Derivation of verbs may change their semantics, either their meaning or their argument structure. Therefore, verbs derived through suffixation can be divided into both intransitive and transitive. The former are found to be formed from nouns in Chorote and Nivacle, whereas the latter are present in all languages of the family. Causatives and applicatives increase the verbal valency, on the one hand, and locatives and directionals change the meaning of the base, on the other hand. I will not go into all locatives and directionals in each language in depth here, but I will rather illustrate the types of derivational processes (the ones that change the verbal valency and the meaning of the base but not the argument structure).

39.3.2.1 *Derived Intransitive Verbs*

Nivacle and Chorote have suffixes to derive intransitive verbs from nouns, thus involving a transposition. The verbalizers in both languages are formally similar. In Nivacle there are two, *-xan* ~ *-xun* “VBZ 1” and *-xɬa* “VBZ2.” According to Stell (1987: 316), *-xɬa* expresses that the thing is taken out from its natural place or environment.

- (58) Nivacle
 a. *-apéx* ‘satisfaction’
 b. *-ape-n-xan* ‘to satisfy’
- (59) Nivacle
 a. *-af* ‘feather’
 b. *-n-fa-xɬa* ‘to pluck’
 (Stell 1987: 316)

On the other hand, Chorote uses the suffix *-xi* “VBZ,” as in (60)–(61):

- (60) Chorote
 a. *-jel* ‘breath’
 b. *a-jeli-xi* ‘I breathe’
- (61) a. *-kjuni* ‘sweat’
 b. *si-kjuni-xi* ‘I am sweaty’
 (Gerzenstein 1978: 122)

To create intransitive denominal verbs, Wichi deploys conversion. Hence, without adding an explicit morpheme, the alienable nominal base changes the verb word-class and can receive subject pronominal prefixes: *nòyij* ‘the way’ > *n’-nòyij* ‘I walk’; *wiyis* ‘blood’ > *n’-wiyis* ‘I am bleeding’; *mamse* ‘young person’ > *n’-mamse* ‘I am a young person.’

39.3.2.2 *Derived Transitive Verbs*

Transitive verb derivation generally consists of valency increasing processes, mainly causativization and applicativization. Causative derivation changes the event structure by adding an agentive participant to the verb. The semantic role of this participant is the “causer” and the syntactic role is the “subject.” Consequently, the subject of the

intransitive verb becomes the object. Mataguayan languages have two different causatives that are formally similar: *-hat* and *-yen* (Cho); *-hit* ~ *-kit* and *-(in)en* (Ma); *-xat* ~ *čat* ~ *-it* and *-(xa)jan* (Ni); *-hat* ~ *-kat* and *-yen* (Wi). The suffix **-hVt* combines with verbal bases that have non-agentive or non-volitional subject:

- (62) Chorote
ityuhmat
 i-tum-hat
 3_{SBJ}-located/to_place_oneself-CAUS
 'He/she puts/places something.'
 (Carol 2009/10: 83)
- (63) Maka
h-iwef-it
 1_{SBJ}-get_tired-CAUS
 'I make somebody get tired.'
 (Gerzenstein 1994: 108)
- (64) Nivacle
-nkafun-xat
 -be_weak/soft-CAUS
 'to make somebody/something weak'
 (Stell 1987: 312)
- (65) Wichi
ihwokw icho-hot *hal'o lhacha*
 wind 3_{SBJ}:to_break(intr)-CAUS tree branch
 'The wind broke the branch of the tree.'
 (Nercesian 2011a: 338)

The other causative, **-jen* ~ *-yen* ~ *-inen*, combines with agentive and non-agentive verbs. The degree of volition and manipulation of the causer is higher than that added by **-hVt*.

- (66) Chorote
hwan i-lyakihn-yen *na pedro*
 John 3_{SBJ}-to_dance-CAUS dem Peter
 'John makes Peter dance.'
 (Carol 2009/10: 81)
- (67) Maka
hi-wet-inen *n-a'* *y-as*
 1_{SBJ}-to_climb-CAUS dem-masc 1_{POSS}-son
 'I make my son climb.'
 (Gerzenstein 1994: 104)

- (68) Nivacle
-ap-xajan
 -to_cry-CAUS
 ‘to make somebody cry.’
 (Stell 1987: 313)
- (69) Wichi
ikatah-yen-n’u
 3SBJ:to_cook-CAUS-1OBJ
 ‘He/she makes me cook.’
 (Nercesian 2011a: 335)

Applicatives also change the event structure of the verb, but they add an object rather than a subject, with the semantic role that the applicative gives. In Mataguanan languages, locatives, directionals, and instrument function as applicatives depending on the verb with which they combine.

- (70) Wichi
- a. *n’-t’on-’am-che*
 1SBJ-to_shout(intr)-2OBJ-APPL.in_extension
 ‘I shout to you (you are going away from me).’
 - b. *n’-t’on-’am=pe’*
 1SBJ-to_shout(intr)-2OBJ=APPL.on
 ‘I outshout you.’
 - c. *n’-t’on-’am-kwe*
 1SBJ-to_shout(intr)-2OBJ-APPL.allative
 ‘I cry out for you.’
- (Nercesian 2011a: 342)

Causatives and applicatives are extremely productive derivational processes on verbs and a noteworthy feature of the family and the whole Chaco area. Applicativization in Wichi shows some morphological differences with respect to causativization (Nercesian and Vidal, forthcoming). On the one hand, the applicatives’ behavior does not change the morphological class of the verb encoded by the prefix *t(V)-* (in monovalent stems) and *(y)i-* (in bivalent ones) as is the case with causativization. Nonetheless, the verbal base changes its valency and allows an applied object. On the other hand, the applicative fills the last slot in the morphological structure of the verb, in contrast with the causative, which is next to the root. Compare example (69)—with the causative before the object: Root+CAUS+OBJ—with (70)—where the applicative is at the end: Root+OBJ+APPL. These differences, along with others of a phonological

nature, are proof that applicativization is a partially morphological and partially syntactic process, or, at least, that it is a morphological process that preserves syntactic traces.

Lastly, locatives and directionals modify the meaning of the verb but not the event structure, as attested in Wichi (71)–(74).

- (71) a. *yukwaj* ‘(s)he bites’
 b. *yukwaj=hi* ‘(s)he chews’
 [₃SBJ:to_bite=LOC.in]
 (72) a. *yihon* ‘(s)he follows something’
 b. *yihon-a* ‘(s)he takes with him/herself, uses’
 [₃SBJ:to_follow=LOC.near]
 (73) a. *tiyoj* ‘(s)he throws, shoots’
 b. *tiyoj=pho* ‘(s)he jumps up’
 [₃SBJ:to_throw=DIR.up]
 (74) a. *tun=hi* ‘(s)he stretches something’
 [₃SBJ:to_take_out?=LOC.in]
 b. *tun=pho* ‘(s)he raises something’
 [₃SBJ:to_take_out?=DIR.up]

(Nercesian 2011a: 371)

In (70a and 70b) the verbal root *tun-* and the locative/directional suffix are colexicalized. Thus, synchronically, *tun-* never appears on its own without the suffix, making it difficult to ascertain the meaning of this verbal root.

39.3.2.3 Multiple Derivational Suffixation

Multiple derivational suffixation in verbal bases is less frequent than in nouns. The only cases with more than one derivative I have documented in Wichi conflate the reflexive with the applicative *-hu*. No causative+causative, causative+applicative, or applicative+applicative combinations are permitted on the same base.

- (75) *n’-lhi-katay-hu*
 ₁SBJ-RFL-to_COOK-APPL
 ‘I cook for myself.’
 (Nercesian 2011a: 329)

The reflexive prefix combines with transitive bases; therefore, the intransitive verb *-katay* ‘to cook’ must be first converted to transitive, adding *-hu* ‘APPL,’ to allow the reflexive *lhi-* and to express the co-referentiality between subject and object. Conversely, reflexive derived verbs allow the applicative to increase the valency and receive an object:

- (76) *n'lhichufwajhu fwa'ay*
 n'-lhi-chufwaj-hu fwa'a-y
 1SBJ-RFL-grind-APPL carob-PL
 'I grind carob (*Prosopis alba*) for myself.'

In (76) the transitive verb *-chufwaj* 'grind' is converted to intransitive as a result of adding the reflexive *lhi-* and then converted to transitive again to allow the object *fwa'ay* 'carob.'

39.4 CONCLUSIONS

To summarize, across Mataguayan languages the main derivational process in terms of productivity and diffusion is suffixation. The languages show uniformity concerning this feature. Nouns and verbs are the word-classes likely to be derived and the variety of derivational processes is wider in nouns than verbs. "Agent," "patient," "action," "instrument," "locative," "evaluative," and "vegetable class" are possible nominal derivations; whereas verbs can be formed by a "verbalizer," "causative," and "applicative." Nominalization and verbalization change the syntactic word-class, and agent, patient, and action can change the nominal class: alienable/inalienable.

Polysemy is not a prominent feature in derivational suffixes. A notable overlap of "agent" and "instrument" is related to the animacy of the referent and the different suffixes that have developed in each language. Also interesting is the evolving of the diminutives from the word "son," and the suffixes that derive nouns as they sort vegetables into fruits, wooden plants, and communities. In addition, more than one derivational suffix is allowed on the same nominal or verbal base.

This analysis provides a preliminary overview of the main derivational processes of the family in nouns and verbs. I have focused on Wichi when the data of the other languages is insufficient, at least for now. Yet, we expect that more in-depth research of the other languages will enable an enhanced the study of the morphology of this family.

CHAPTER 40

AREAL TENDENCIES IN DERIVATION

BERND HEINE

40.1 INTRODUCTION

THE present chapter is concerned with how derivational structures may arise in situations of language contact, giving rise to the areal spread of grammatical features or categories.

There are two main linguistic mechanisms that can lead to areal spread. One involves *borrowing*, that is, the transfer of form–meaning units from one language to another; in the model of Johanson (1992, 2002a, b) this is referred to as global code copying. The history of English offers a plethora of examples showing that the borrowing of derivational forms provides a fairly common and convenient strategy for acquiring new grammatical concepts and categories (see e.g. Miller 2006). A few examples of English borrowings from Greek may suffice to give an impression of the magnitude of the process concerned:¹

- (1) Examples of English derivational prefixes borrowed from Greek.

English	Greek source
<i>auto-</i>	<i>autos</i> ‘self’
<i>exo-</i>	<i>exo</i> ‘outside’
<i>homo-</i>	<i>homos</i> ‘sane’
<i>mono-</i>	<i>monos</i> ‘single, alone’
<i>neo-</i>	<i>neos</i> ‘new’
<i>syn-</i>	<i>syn</i> ‘together, same’
<i>tele-</i>	<i>tele</i> ‘far’

¹ Note that most of these borrowings are of limited derivational productivity.

Similar borrowings from Greek and Latin have also taken place in a number of other European languages, to the extent that there is some justification to consider them an areal feature of Europe (cf. Haspelmath 1998, 2001, Heine and Kuteva 2006).

Borrowings of the kind illustrated in (1) have been well described in English and a number of other languages. One reason for why such borrowed derivational forms have attracted a remarkable interest among students of language is presumably because they are usually not hard to detect: The fact that there is similarity or even identity in both form and meaning between the source and the target languages makes it, as a rule, easy to identify them and to reconstruct the borrowing processes concerned.

It is much more difficult to detect cases where the mechanism involved is not borrowing but rather *replication*, that is, where meanings and/or structures but no forms are transferred from one language to another. Accordingly, there is much less information on this mechanism, especially on what replication means with reference to the areal distribution of derivational forms and structures. In the remainder of this chapter I will be restricted to grammatical replication, summarizing the few findings that have been made on the contact-induced transfer of derivational and related grammatical patterns.

Derivation as a morphological domain is not a characteristic of human languages in general; it is in particular largely or entirely absent in languages of the analytic-isolating type, such as the languages of the West African coastal belt or of Mainland South East Asia. But grammatical functions frequently expressed by means of derivational mechanisms, such as causativity on verbs, can be encoded as well by lexical means, not only in internal language development but also in language contact, as we will see in Section 40.3.

40.2 GRAMMATICAL REPLICATION

Following Heine and Kuteva (2003, 2005, 2006) the terms *model language* (in short: M) and *replica language* (R) are used for the languages being, respectively, the source (or “donor”) and the target (or “recipient”) of transfer. These terms are adopted from Weinreich ([1953] 1964: 30–1), and so is the term replication, which stands for kinds of transfer that are traditionally referred to as “calquing,” “structural borrowing,” and the like (e.g. Ross 2007). Note that there are many alternative terminologies. Thomason and Kaufman (1988) or Thomason (2001: 93), for example, use borrowing for both kinds of transfer. Furthermore, I will distinguish between replication restricted to the lexicon (= lexical replication) and replication that concerns grammatical meanings or structures (= grammatical replication). In accordance with the topic of this volume, my concern is exclusively with the latter, that is, with a process whereby speakers create a new grammatical meaning or structure in language R on the model of language M by using the linguistic resources available in R. Note that R may be either the mother tongue (L1) or a second language (L2) of the speakers concerned, and the same applies to M.

As has been shown in Heine and Kuteva (2003, 2005, 2006), grammatical replication is almost invariably in accordance with principles of grammaticalization, and so are the data that I will be concerned with here.² Contact-induced grammaticalization is a grammaticalization process that is due to the influence of one language on another; instances of it have been discussed in a wide range of studies (see especially Campbell 1987, Stolz 1991, Haase 1992, Heine 1994, Harris and Campbell 1995, Nau 1995, Aikhenvald 1996, 2002, 2003, 2006, Bisang 1996, Kuteva 1998, 2000, Dahl 2000, Stolz and Stolz 2001, Heine and Kuteva 2003, 2005, 2006; see also the contributions in Aikhenvald and Dixon 2001, 2006), even if a number of these studies do not use a framework of grammaticalization theory.

40.3 FROM COMPOUNDING TOWARDS DERIVATION

Not uncommonly, context extension involves a grammaticalization process whereby the use of an item belonging to some open morpheme class is extended to serve as a modifier of another morphological paradigm. In this way, nouns and verbs may lose their categorial status and become clitics or affixes of other nouns or verbs. The following Finnish example, taken from Nau (1995: 86–7), may illustrate the process concerned. Presumably on the model of English *top* (and Swedish *topp*), speakers of Finnish have extended the use of their noun *huippu* ‘top, peak’ to act as a modifier of nouns in compounds, which thereby assumed the schematic meaning ‘an outstanding, a very important X.’ Some examples are provided in (2a). This use pattern is found especially in journalistic writing, but it is not the only one of its kind. Nau describes a second one, involving the Finnish noun *avain* ‘key’ which has also given rise to a compounding pattern on the model of English ‘the key X’; examples are provided in (2b).

(2) Finnish (Nau 1995: 86–7)

a.	<i>huippuhenkilö</i>	‘an outstanding personality’
	<i>huippuurheilija</i>	‘a top sportsman’
	<i>huipputekniikka</i>	‘top technology’
b.	<i>avainasema</i>	‘key position’
	<i>avainhahmo</i>	‘key figure’
	<i>avainkysymys</i>	‘key question’

² Grammaticalization is defined in terms of the following four parameters: (a) *extension*, i.e. the rise of novel grammatical meanings when linguistic expressions are extended to new contexts (context-induced reinterpretation), (b) *desemanticization* (or “semantic bleaching”), i.e. loss (or generalization) in meaning content, (c) *decategorialization*, i.e. loss in morphosyntactic properties characteristic of lexical or other less grammaticalized forms, and (d) *erosion* (or “phonetic reduction”), i.e. loss in phonetic substance (Heine and Kuteva 2005: 15).

We seem to be dealing with the emergence of two contact-induced use patterns in Finnish, though neither of them is suggestive of a major use pattern, both being of limited productivity. The two have been created on the same model and based on the same strategy, where the first of two nouns in compounds is grammaticalized to a modifying constituent, being extended to a specific context (= extension), losing its specific lexical semantics in favor of some schematic function (= desemanticization), and losing nominal properties, such as the ability to be modified or inflected (= decategorialization; concerning these parameters, see Section 40.2). And both grammaticalizations are suggestive of an incipient process leading towards derivation.

Derivational morphology can be historically derived from a wide range of lexical forms, most of all from nouns and verbs. But adverbs and adpositions are also among the forms leading to the rise of derivational elements, both in internal grammatical development and in language contact. Kostov (1973: 110–13) discusses one common type of what he calls “interference” phenomena under bilingualism where the receiving language develops a pattern that is only occasionally used into a regular pattern under the influence of another language. As an example he cites the Loväri variety of Romani whose speakers use Hungarian as a second mother tongue and who grammaticalized their own locative adverbs to derivational³ verbal prefixes in an attempt to replicate Hungarian verbal prefixes, thereby introducing a new morphosyntactic pattern for which earlier there was essentially no equivalent in Loväri.⁴ Note that verbal composition was not unknown to Loväri speakers; what they did under Hungarian influence was simply increase the frequency of use of a rare type of expression, and to extend and elaborate it. This strategy had quite some impact on the morphosyntactic profile of the Loväri variety (Kostov 1973: 113).

A similar example is provided by the contact situation between the genetically unrelated Bantu and Nilotic languages of western Kenya (Dimmendaal 2001), where Bantu languages appear to have acted as model languages. Bantu languages are known for their rich paradigms of verbal derivational extensions marked by suffixes. There is nothing comparable in the Western Nilotic language Luo or its closest relatives, the Southern Lwoo languages of Uganda and the Sudan: Verbal derivation is limited, mainly involving internal morphology in the verb root. Luo has been in contact with various Bantu languages for nearly five hundred years. Apparently on the model of neighboring Bantu languages, Luo speakers have developed a set of what look like verbal derivational suffixes, resembling structurally the Bantu verbal suffixes and expressing functions typically encoded by the Bantu derivational applied suffix **-id-* (‘for, to, with reference to,

³ Kostov (1973: 110–11) refers to the process as verbal compounding rather than derivation. Unfortunately, Kostov does not provide any unambiguous examples to illustrate this process.

⁴ “Aber die ungarische Sprache, deren sich die Loväri-Zigeuner in Ungarn als zweiter Muttersprache bedienen, hat dazu beigetragen, daß die Loväri-Mundart über ein herausgebildetes, völlig selbständiges System der Verbalkomposition verfügt” [But the Hungarian language, used by the Loväri gypsies in Hungary as a second mother tongue, has contributed to the fact that the Loväri dialect disposes of an elaborate and completely independent system of verbal compounding] (Kostov 1973: 111).

on behalf of'). The process involved appears to have been grammaticalization, whereby adpositions lose their independent status and turn into verbal derivational clitics and affixes. Luo speakers used their prepositions *ne* (or *nI*) benefactive, *e* locative, and *gi* instrumental to develop verbal enclitics or suffixes;⁵ the following example is confined to the benefactive preposition *ne*. Example (3a) illustrates the prepositional use and (3b), where *Juma* is topicalized, the use as a verbal suffix.

(3) Luo (Western Nilotic, Nilo-Saharan; Heine and Reh 1984: 51)

- a. *jon nego diel ne juma.*
 John is.killing goat for Juma
 'John is killing a goat for Juma.'
- b. *juma jon nego- ne diel.*
 Juma John is.killing- for goat
 'John is killing a goat for Juma.'

That in situations of language contact speakers draw on existing lexical material to create patterns of derivation that they find in the model language can also be illustrated with the following example from Basque. For roughly two millennia, Basque has been in close interaction with a range of Romance languages, most of all Latin, Spanish, Gascon, and French, and this interaction has left its marks on the structure of the language. For example, derivational prefixing on verbs is something to be found in Romance languages but very rarely in traditional Basque. There is evidence to suggest that Basque is moving in the direction of Romance by developing an equivalent for the Romance [prefix—verb stem] construction. For example, Jendraschek (2006: 158–9) found that the verbal derivational prefix *re-* is frequent in Romance, and Basque speakers have developed an equivalent of this derivational element by using the form *bir-/berr-*, possibly a grammaticalized form of *berri* 'new', as an equivalent. As the examples in (4) suggest, this form is both structurally and semantically a fairly direct equivalent of the Romance prefix.

(4) Examples of the Basque verbal prefix *bir-/berr-* (Jendraschek 2006: 158)

<i>aztertu</i>	'examine'	<i>berr-aztertu</i>	're-examine'
<i>erabili</i>	'use'	<i>berr-erabili</i>	'reuse'
<i>sortu</i>	'create,' 'found'	<i>bir-sortu</i>	'regenerate'

As observed earlier, it is not always easy, if at all possible, to trace a clear boundary between derivational and inflectional structures. For example, passives and case marking devices tend to be associated with inflection, but they may also share features to be expected in derivation. The present discussion therefore includes an example from an erstwhile Portuguese-based creole of Sri Lanka where language contact appears to have been responsible for the grammaticalization of case affixes.

⁵ The Luo elements concerned have features of both suffixes and enclitics.

Over three hundred years of language contact in Sri Lanka between the Dravidian language Tamil on the one hand and Malay and Portuguese on the other had the effect that the latter two, nowadays referred to as Modern Sri Lanka Malay and Modern Sri Lanka Portuguese, respectively, developed into creoles with basic verb-medial (SVO) word order. More recently, the two creoles replicated grammatical structures of Tamil, to the extent that they turned into verb-final languages like Tamil, developed roughly the same case system, evidentials, quotative markers, semantics of verb markers, etc.—just like the model language Tamil (Bakker 2000). Contact-induced grammaticalization played an important role in the Tamilization process.

Neither of the two creoles had a grammaticalized case system. In an attempt to replicate the case system of Tamil, speakers of the creoles used a universal process of grammaticalization in developing adpositional, adverbial, and other independent forms into case suffixes, using the lexical resources that were available in the creoles. Thus, the Portuguese adposition *pera/para* ‘for’ was grammaticalized to an accusative/dative case suffix *-p<->*, Portuguese *junto* ‘together’ to a locative case suffix *-(u)nto*, or the Portuguese possessive pronouns *sua, seu* ‘his, her, its’ to a genitive suffix *-su(wə)* in Modern Sri Lanka Portuguese, thereby matching the structure of the corresponding Tamil case markers (Bakker 2000: 32, Heine and Kuteva 2010).

Example (5) below relates only marginally to derivation but it may illustrate the development of an areal dimension in contact-induced language change.

Passive markers in a number of European languages commonly use “be” or “become” as auxiliaries, with the main verb being encoded as a perfect participle form or some equivalent of it. But, as Ramat (1998b) shows, Rhaeto-Romance, Italian, and the Bavarian dialect of German share a periphrastic passive construction based on the grammaticalization of the lexical verb “come” to a passive auxiliary, cf. (5).

(5) The Alpine “come”-passive (Ramat 1998b: 227–8)

a. Ladin (Rhaeto-Romance)

Cô vain fabricheda la scuola nouva.
 (here comes built the school new)⁶
 ‘Here the new school is constructed.’

b. Italian

Qui viene costruita la scuola nuova.
 (here comes built the school new)
 ‘Here the new school is being constructed.’

c. Bavarian (German)

Dâ kummt de nei(e) Schul gebaut.
 (here comes the new school built)
 ‘Here the new school is constructed.’

⁶ Ramat (1998b) does not provide interlinear glosses for these examples; the use of parentheses indicates that the glosses are ours.

This kind of areal patterning, based on one and the same process undergone by neighboring languages, is called a grammaticalization area by Heine and Kuteva (2005: 182–6). That the Alpine “come”-passive in fact forms a grammaticalization area is suggested by the following observations: The constructions in the three languages concerned are the result of the same general process of grammaticalization from a construction [“come” + perfect participle] to a passive construction, where a lexical verb for “come” gave rise to a passive auxiliary. And this process must have involved language contact, for the following reasons: First, “come”-passives do not appear to be cross-linguistically common and it is therefore statistically unlikely that these three neighboring languages should have undergone such a rare process independent of one another.⁷ Second, common genetic inheritance can be ruled out as a contributing factor: The “come”-passive can be traced back to earlier stages of neither Romance nor Germanic languages. And third, the languages concerned can be shown to share a history of contact; Ramat (1998b: 227–8) notes that there have been centuries of contact between southern Germany and northeastern Italy and that it was the Romance type that “has influenced the geographically contiguous Bavarian passive.” Accordingly, areal diffusion of a grammaticalization process among these geographically adjacent communities offers the most reasonable explanation.

Among the grammatical concepts that turn out to be particularly dynamic in language contact and areal spread, there is that of causativity: Deriving causative expressions from non-causative verbs, but also from nouns or adjectives, is a process that can be shown to be induced or accelerated by language contact. In situations of language contact, it is most of all verbs meaning “do, make” or “leave” that tend to grammaticalize into causative markers. As a rule, this process does not lead to the rise of new affixal causative markers but tends to be restricted to lexical expressions for causative concepts.

The following example may illustrate this process. The example is taken from Basque, more specifically from language contact in France, where Gascon and French provided the main model languages. As Haase (1992: 125–6) argues convincingly, it was either of these two languages that must have provided the model for the grammaticalization of a causative construction in Basque as spoken in France: Gascon *har* ‘do’ + infinitive and French *faire* ‘do’ + infinitive are both used as lexical causatives, and either of them appears to have been responsible for the fact that the Basque verb *egin* ‘to do’ grammaticalized to some extent into a causative marker, used in a participial form, as in the following example:

(6) Basque (Haase 1992: 125–6)

[...] <i>ber-</i>	<i>e-</i>	<i>k</i>	<i>lei[h]o</i>	<i>bat-</i>	<i>eta-</i>	<i>(r)at</i>
IDN-	ART-	ERG	window	one-	TRN-	DIRV
<i>jautz-</i>	<i>i</i>	<i>[eg]in</i>	<i>eta [...]</i>			
descend-	PTCP	do.PTCP	and			
‘[...] they must take them (down) to a window and [...]’						

⁷ The only other case we are aware of is found in Maltese.

Similar examples are not hard to come by. For instance, in the language contact situation of Amazonia in northwest Brazil, the North Arawak language Tariana appears to have replicated a lexical causative from East Tucanoan languages, involving the grammaticalization of the verb *-ni* 'do' (Aikhenvald (2002: 116). Causatives do not only play a role in the expression of causative meanings but may also serve other grammatical purposes. In Northern Chinese there are two passive constructions formed with the causative verbs *jào* 'make' and *ràng* 'let', and Norman (1982: 245) argues that the growth of these constructions was not an independent development but rather was due to superstrate influence by the Altaic language Manchu on Chinese (see Chappell 2001: 350).

The transfer of causative concepts is by no means restricted to lexical processes. There are a number of examples to show that quite commonly, morphology is involved in a contact-induced transfer process, and it can concern both grammatical replication and borrowing.

Thomason (2001: 112) observes that in most Semitic languages, causatives are expressed by means of single prefixes whereas Ethiopic Semitic languages use a double prefix. She hypothesizes that the latter structure has been modeled on the double-suffix formation of Cushitic (substratum) languages. In addition to grammatical replication there is also evidence for the borrowing of derivational morphology. For example, in the contact situation of the Sepik River basin of northern Papua New Guinea, the Yimas language appears to have borrowed from Enga the causative suffix *-(a)sa*, added to a small set of verbs (Thomason 2001: 119–20).

40.4 EXTENSION ACROSS WORD CATEGORIES

But the process leading to the development of new derivational forms is not restricted to lexical items; it may as well involve the extension of a grammatical element from one function to another. The more a given use pattern expands, the more likely it is that the pattern will involve new grammatical structures. The moribund Aztec language Pipil of El Salvador has been strongly influenced by Spanish and, as a result, has experienced a number of structural changes (Campbell 1987, Harris and Campbell 1995). For example, Pipil has a category of past participial forms, marked by *-tuk*, which formerly functioned only in verbal constructions, as in (7a). Apparently, Pipil speakers established an equivalence relation between their category and the Spanish participle, which functions not only in verb forms (e.g. *ha roto* 'has broken') but also often in adjectives (*el vidrio roto* 'the broken glass'). On the basis of the model provided by Spanish, the use of the Pipil past participial forms has been extended to new contexts, "to bear an adjectival function as well," as Campbell puts it (glossed PERF by Campbell 1987: 265), as in (7b).

(7) Pipil (Aztec, Uto-Aztec, Campbell 1987: 265)

a. *ki-chiw-tuk*.

it- do- PAST.PARTICIPIAL

'S/he has done it.'

- b. *kabál uksi- tuk ne wahkal.*
 exact ripe- PERF the gourd
 'The gourd was just ripe.'

Another pathway of contact-induced extension concerns a process whereby the use of markers associated with the noun phrase is extended to also mark clause subordination. In this way, demonstrative attributes commonly develop into relative clause pronouns, case and nominalization markers into markers for complement clauses and adverbial clauses, etc. (see Heine and Kuteva 2002, 2005 for examples). The North Arawak language Tariana of northwest Brazil provides an example for such an extension pattern. Tariana speakers form complement clauses (that is, clauses that occupy a core argument slot of the predicate) by means of the subordinator *-ka*. But they have been in close contact with speakers of East Tucanoan languages and as a result have replicated many use patterns of these languages. East Tucanoan languages use nominalizations for introducing complement clauses, and younger and innovative Tariana speakers tend to replicate this structure by extending their nominalization markers to also introduce complement clauses on the model of East Tucanoan languages (Aikhenvald 2002: 162). Note that this is an on-going process that so far has not spread throughout the entire Tariana-speaking community.

40.5 NEW FORMS OF NOUN CLASSIFICATION

Another grammatical domain that tends to be located at the interface of derivation and inflection is that of noun class systems. In languages having such a system, noun classes belong to the parts of grammar that are of frequent occurrence in discourse, and they as well may be affected by language contact. There is at least some evidence to suggest that such systems can be partly or fully replicated in situations of language contact via grammaticalization. As Dixon (2002: 500–1) and Aikhenvald (2006: 18) found, for example, the Yanyuwa language of Australia developed a system of noun classes via grammatical replication in imitation of the system of its westerly neighbors.

Further examples are provided by a couple of languages in Amazonia. The linguistic area of the Vaupés basin in northwestern Brazil provides a plethora of examples where some languages, such as the North Arawak language Tariana (Aikhenvald 2002, 2006) or the Nadahup (Makú) language Hup (Epps 2006), have been massively influenced by East Tucanoan languages in general and Tucano in particular, and where grammatical categories of the replica languages Tariana and Hup exhibit only a weak degree of grammaticalization, those of the Tucanoan model languages are distinctly more strongly grammaticalized. For example, the Tucanoan languages have a system of noun classification where inanimate items are categorized by shape and animate ones by gender, and Hup speakers are in a process of developing a totally

new system of classifiers under the influence of Tucano (Epps 2006, Aikhenvald 2006: 16), grammaticalizing for example nouns for plant parts into shape-based classifiers on the model of Tucanoan languages. As is to be expected in language contact, the Hup classifier system is distinctly less grammaticalized; for example, unlike its Tucanoan model, it is not generalized but rather is restricted to only a small subset of nouns.

40.6 CONCLUSIONS

Borrowing (or global copying) of derivation is well-documented and appears to be a widespread phenomenon (Thomason and Kaufman 1988, Thomason 2001). The main goal of the present chapter was to show that derivational forms or features can also be *replicated* via the transfer of structures or grammatical meanings from one language to another. This process has so far not received the kind of attention in scholarly work that it deserves. As the presentation may have shown, we still lack sufficient data to understand the mechanism in every detail.

The main observation to be made on the contact-induced, or areally induced growth of derivational forms and structures involves a process of the following kind: Speakers draw on the linguistic resources that are available in the replica language, which may be either their mother tongue (L1) or a second language (L2), to develop a new grammatical category, such as a derivational form, on the model of another language providing the model for replication. The process involved is usually, though not necessarily, contact-induced grammaticalization as described in Heine and Kuteva (2005, 2006). This process is gradual and may take centuries to be completed. In fact, in quite a number of the examples that we surveyed above, we were confronted with incipient stages of the process rather than full-fledged new categories. The main result of such processes is that the languages involved acquire a higher degree of translational equivalence and areal relationship.

CHAPTER 41

UNIVERSALS IN DERIVATION

ROCHELLE LIEBER AND PAVOL ŠTEKAUER

41.1 INTRODUCTION

TYPOLOGICAL research into language structures, including research into language universals, has made considerable advancements in last few decades, as attested by the increasingly frequent appearance of general and pedagogical works like Whaley (1997), Song (2011), Velupillai (2012), and Pereltsvaig (2013). But the vast bulk of research has concentrated on issues other than derivation. For example, in Song (2011), there is only one chapter on morphology (Brown 2011), and it is primarily devoted to a discussion of inflection. Although the present volume has not been intended as a contribution to the field of typology, the sheer concentration of material we have gathered here can aid us in filling some of the gaps, and perhaps in assessing the issue of universals in derivation. We begin in Section 41.2 by discussing some of the obstacles to looking for universals of derivation. Section 41.3 reviews some of the past proposals for universals of derivation. Finally, in Section 41.4 we assess some of these proposed universals in light of the material assembled in this volume.

41.2 METHODOLOGICAL PROBLEMS

In looking at derivation from a typological perspective and seeking to find universals, we are faced with a number of problems from the outset. Some of these are general problems in the field of typology and others are specific to word formation.

41.2.1 Sampling

Our first methodological issue is not unique to typological work on derivation. The bane of all typological work is the issue of sampling: what constitutes a representative sample

of languages on which we might base putative universals? This is an issue that has been prominent in the literature on typology in general; see for example, the overview by Bakker (2011). Under ideal conditions, a sample would be large and well-balanced with regard to geographical and genetic coverage including language isolates. But conditions are never ideal. An ideal sample presupposes the availability of high quality descriptive grammars for all the languages we might wish to sample. Further these grammars must contain material that is not only relevant and comprehensive, but also comparable across languages. In reality, the necessary data are rarely available. The quality, depth, emphasis, and theoretical orientation of available descriptive grammars varies widely, and unfortunately, the attention to word-formation (as opposed to inflection) in such grammars is frequently rather superficial.

The alternative to basing typological work on descriptive grammars is to obtain data from informants, but this method too is fraught with difficulties, foremost obviously being the problem of finding knowledgeable informants. Indeed, for some languages, especially endangered ones, there may be just one or a small handful of experts who may or may not have given much attention to derivational morphology.

41.2.2 Terminology

Also not unique to the typological study of derivation is the issue of terminology, but it is a problem that is especially acute with respect to morphology. If a researcher is to draw generalizations based on a large sample of data, the data should ideally be organized according to the same well-defined concepts. Unfortunately, the study of word-formation is perennially burdened with the problems of defining even its most basic units: morpheme, word, root, stem, and so on. Theorists may use alternative terms for what might or might not be the same process cross-linguistically (for example, conversion or zero affixation). Or there may be no standard term at all for a concept: terms like “linking element,” “interfix,” “stem extender” for the element that may occur between words or roots in compounds come to mind. Or different theorists might include different phenomena under a given rubric, as Davis and Tsujimura (this volume) point out with respect to non-concatenative morphology. If linguists are working with different notions of what constitutes a word, a stem, or an affix, our task in comparing derivation in the languages of the world is vastly more complex: the way we understand the nature of basic units may determine the classification of a word-formation process that underlies the creation of a complex word—an issue crucial to research into typology and universals.

41.2.3 The Influence of Theoretical Analysis

Problems can also result when several potentially different theoretical analyses are available for a particular phenomenon. Consider the question of whether back-formation

should be treated as a matter of derivation. There are a number of ways of looking at back-formation, depending on one's theoretical perspective, as Štekauer (forthcoming b) illustrates. Thus, one group of authors admit only the diachronic relevance of this process (for example, Marchand 1963b); if back-formation is a purely diachronic process, it is not something we should concern ourselves with in a synchronic typological study of derivation. On the other hand, there are linguists who stress that, synchronically, back-formation is a productive word-formation process (in English), no matter what its nature stems from. In this group, Kiparsky (1982b), *inter alia*, conceives of "back-formation" of the *air-condition*_V type as a process of compounding. In contrast, Nagano (2007) maintains that back-formation is best accounted for as conversion with subsequent clipping. The process that gives rise to words like *baby-sit*, *brainwash*, *housebreak*, *book-keep*, *spring-clean*, *sight-read*, *strap-hang*, *sleep-walk*, *chain-smoke*, *vacuum-clean*, etc. might also be accounted for as noun-incorporation as proposed by Hall (1956), Brömser (1985), and Štekauer (forthcoming b), among others. Each of these approaches results in a different typological perspective.

41.2.4 The Problem of Boundaries

Chapters 2 and 3 have already surveyed the issues of determining the boundaries between derivation with respect to inflection on the one hand and compounding on the other. Obviously, typological comparisons can be affected by the decision as to where we set these boundaries. Further, regardless of where we set the boundary, the same category may be inflectional in one language and derivational in another. A case in point is evaluative morphology. It is well known, for example, that while evaluative morphology looks derivational in many languages, it looks more like inflection in some languages, for example, several of the Niger-Congo languages.

At the heart of derivation we look at means for creating new lexemes in languages. Word-formation is concerned with the creation of new complex words, but the term "new word" is vague. It may be defined as a name given to an object (in the widest sense of the word) of extra-linguistic reality by (predominantly) morphological techniques or strategies. But words may be new in different ways. Most clearly, a name might be given to an object which has not been labeled by a linguistic sign yet. But beyond this, a new linguistic sign might be formed to denote an existing object based on a more relevant and/or updated motivation and/or to reflect some other shades of the object in question. It is the latter option that gives rise to problems.

For example, do clippings and acronyms (as mere form-reductions) comply with the aforementioned definitions? Certainly they are not new in the first sense, but do they fit the second definition enough to be counted as new words? And if not, should they not be studied as part of derivation? Similarly, we might look at the status of evaluatives with recursively used evaluative markers. Consider Sl. *psíček* 'very little or much loved dog,' which is derived from *psíča* 'little or well-loved dog' which in turn is derived from *pes* 'dog' with double diminutivization by two different evaluative markers. Is this to be

considered a new word, different from *psíča*, especially as they belong to different paradigms? And what about the form cited by Johns (this volume), in which the same evaluative marker is used recursively in Labrador Inuktitut?

- (1) *Kuatsâ-pâ-pâ-pâ-pâ-pâ-pâ-lauk-Kunga*
startle-very-very-very-very-very-very-d.PAST-INDIC.1S
'I really really had a very big fright'

This issue is important because if we consider these evaluatives as new words then it might also be argued that the comparative and the superlative degrees of adjectives should also be treated as derivational forms, although they are mostly discussed within inflectional morphology. We are clearly on a slippery slope here.

41.2.5 Determining Productivity

In principle, we would like to base our typological comparisons on word-formation processes or rules that are productive. In practice, there are a number of reasons why it is difficult to implement this simple and logical postulate.

First and foremost, the concept of productivity is scalar, so determining the boundary between productive and unproductive processes is often difficult, especially in languages where documentation of word-formation processes may be sparse. Further, there is no unanimously accepted way of measuring productivity. Various ways of computing productivity have been proposed by Aronoff (1976), Baayen (1989, 1992), and ŠtekaUER (1998, 2005), among others. But even if one method were to be generally selected as a common method of computing productivity, the enterprise would be nearly impossible. Computations of this sort require the availability and detailed study of corpora and complex calculations for every language to be studied, something that is clearly not possible for any but the best documented languages. In the end, we are left to rely on general intuitions of descriptive grammarians and/or informants.

41.3 SOME PAST APPROACHES

In this section, we survey various proposals that have been made in the literature concerning universals or universal tendencies in derivation. Some of the earliest proposals for universals in word-formation can be found in Greenberg (1963a), for example, his Universals 26, 28, and 29:

- (2) a. Greenberg Universal 26: If a language has discontinuous affixes, it always has either prefixing or suffixing or both.

- b. Greenberg Universal 28: If both derivation and inflection follow the root, or they both precede the root, the derivation is always between the root and the inflection.
- c. Greenberg Universal 29: If a language has inflection, it always has derivation.

In more recent work, we have the material available in the Universal Archives (<<http://typo.uni-konstanz.de/archive/intro/>>) in which Frans Plank has gathered 2,029 universals of various kinds. But in this database rather little attention has been paid to derivation specifically. Out of the 2,029 universals included in the Archive, only fifty-one are directly categorized as universals pertaining to word-formation, but oddly most of these actually bear on the word class of numerals rather than on word-formation. Only five seem to us to bear directly on derivation. To these, we can add a few more proposed universals discovered by searching directly for “derivational morphology” in the database, but in the end, only nine of all the assembled universals can be interpreted (somewhat generously) as focusing on synchronic word formation either in isolation or in comparison to inflection:¹

- (3) a. *Univ. 366*: If morphology is agglutinative, then derivational affixes tend to have more complex syllabic structure. If morphology is flexive, then derivational affixes tend to have less complex syllabic structure.
- b. *Univ. 372*: If morphology is agglutinative, then syllabic structure is simple (syllables are short), and if morphology is flexive, then syllabic structure is complex (syllables are long), and vice versa.
- c. *Univ. 605*: If there is a derivational morpheme whose distributional characterization makes reference to objects, then it will also make reference to intransitive subjects but not to transitive ones.
- d. *Univ. 1198*: If morphological categories are marked on verbs rather than nouns, then they specify aspects of events and states rather than of persons and objects, and they are derivational rather than inflectional.
- e. *Univ. 1583*: If there are causative affixes which serve to form causative verbs from transitives, then there will be causative affixes which serve to form causative verbs from intransitives.
- f. *Univ. 1863*: If reduplication is used for grammatical purposes in any other word class, then it is also used (for whatever purpose: gradation, superlative, intensification, distributivity, diminution, . . .) for adjectives or adjective-derived nouns [perhaps rather: property words].

¹ For bibliographical information and classification by kind of universal, see the Universals Archive database.

- g. *Univ. 1939*: If words of whatever word class can be derived from nouns, then nouns can be derived from words/stems other than nouns.
- h. *Univ. 1940*: If there is internal stem modification or suprasegmental modification or subtraction or conversion or suppletion in derivation or inflection, then there is affixation.
- i. *Univ. 2009*: If augmentatives/diminutives can be formed from Determiners, then they can also be formed from Adverbs, Numerals, Pronouns, or Interjections; if from Adverbs, Numerals, Pronouns, or Interjections, then also from Adjectives or Verbs; if from Adjectives or Verbs, then from Nouns.

To these we might add a generalization proposed by Bauer (1997a) and Bakema and Geeraerts (2000), as well as a couple formulated within the theory of Natural Morphology (Dressler 1988):

- (4) If a language has augmentatives, it also has diminutives (Bauer 1997a, Bakema and Geeraerts 2000).
- (5) Dressler (1988)
 - a. Cross-linguistically, word-based rules should be more frequent than root-, stem-, word-form-, phrase-, sentence-based rules.
 - b. Existing (= lexically stored) words can be directly referred to, whereas potential but nonexistent words must first be derived before they can be referred to. Therefore existing words are better bases than potential (but nonexistent) words, and so the occurrence of potential words as bases implies occurrence of existing (actual) words as bases of rules, but not vice versa.

Finally, we can look to recent work by Štekauer et al. (2012) to add to the modest list of possible universals we have compiled above. Using the database of fifty-five languages compiled for that volume, a number of other generalizations pertaining to derivation can be drawn, among them: general universals (6), generalizations pertaining to affixation (7), generalizations pertaining to reduplication (8), and generalizations pertaining to the relationship between form and meaning (9):

- (6) derivation in general
 - a. Every living language has means to coin new linguistic signs.
 - b. Every language makes it possible to form new words by means of concatenative word formation.
 - c. Cross-linguistically, the hierarchy of word-formation processes in terms of their frequency of occurrence is as follows:
 - suffixation, compounding, prefixation
 - conversion
 - back-formation, circumfixation, infixation, etc.

- (7) prefixation, suffixation, infixation
- a. Suffixation is cross-linguistically the most widespread word-formation process (this appears to be true of morphology in general, as pointed out by Sapir 1921, Greenberg 1957, Hawkins and Gilligan 1988, Song 2011).
 - b. If a language uses prefixation for word-formation, it also uses suffixation for word-formation.
 - c. The existence in a language of derivational infixation implies the existence of derivational suffixation.
 - d. The existence in a language of derivational infixation implies the existence of derivational prefixation.
 - e. There is a strong tendency for languages with derivational suffixation to make use of recursive suffixation, i.e., more than one derivational suffix is permitted.
 - f. While suffixal recursiveness is cross-linguistically most characteristic of nouns, prefixal recursiveness is most typical of verbs.
 - g. The existence in a language of multiple derivational prefixation implies the existence of multiple derivational suffixation.
- (8) reduplication
- a. There is a strong cross-linguistic tendency to use reduplication productively for word-formation processes.
 - b. Languages with derivational reduplication show strong preference for full reduplication.
 - c. If there is partial derivational reduplication in a language there is also full reduplication. Exceptions include, for example, Nelemwa (Austronesian), Upper Necaxa Totonac (Totonacan), Movima (isolate), Turkish (Altaic), and Lithuanian (IE).
 - d. In partial reduplications, there is a tendency to copy the stem material on the side from which it is taken, i.e., prefixes tend to copy stem-initial material, suffixes stem-final material, and infixes stem-internal material, although, “relocations” are not infrequent across languages.
- (9) relationship between form and meaning
- a. If a one-to-many relation between form and meaning exists in prefixation of a language, it also exists in suffixation of the same language.
 - b. Instrumental and agent nouns may be derived via the same derivational process, but instrumental and patient nouns are derived by the same process only if agent nouns are also derived by this process.
 - c. There is a tendency for both derivational suffixes and prefixes to feature a many-to-one relation between form and meaning.
 - d. The existence of one-to-many relations between form and meaning of derivational prefixes in a language implies the existence of many-to-one relations between form and meaning of derivational suffixes.

- e. The dominating word-formation process used cross-linguistically to express the semantic category of AGENT is suffixation.
- f. The formation of MALE NAMES ON FEMALE NAMES implies the productive formation of FEMALE NAMES ON MALE NAMES.
- g. The most frequent source of ABSTRACT NOUNS cross-linguistically are adjectives followed by nouns and verbs.

The real question, of course, is how universal such proposed universals are, or alternatively, how strong we find various cross-linguistic tendencies.

41.4 ASSESSMENT

The present volume gives us an unusual opportunity to comment on some of the proposals detailed in Section 41.3. Neither the general chapters in Part 1 of this volume nor the descriptive chapters in Part 2 were solicited with the idea in mind of checking their contents against the various typological generalizations that have been proposed. And not all the proposed universals lend themselves to testing with data gathered in this book; for example, we cannot systematically compare inflection or compounding to derivation here, as authors were asked to write specifically about derivation. Further, some of the proposed universals seem contradictory to us (3a, 3b) or too theory-bound to be able to test (5a, 5b). Nevertheless, the wealth of data discussed in these chapters gives us a rare opportunity to compare at least a few of the proposals to our findings.

What we find is that there is quite a range in the solidity of these various universals. For example, we are fairly certain that the universal in (6a) is an absolute universal: we have not encountered any language or language family in which there is no derivation at all. We are also fairly certain that Greenberg's Universal 28 concerning the ordering of inflection and derivation holds up relatively well, although it was originally based on an analysis of only 30 languages. Here, though there are at least a few known counter-examples in various languages of the world (Nikolaeva, this volume, Bohnert 1984). Further, it is perhaps hard to assess (6a) when we look at languages with templatic derivation, as pointed out in the chapters in this volume on Yenisean (Vajda), Athabaskan (Rice), and also in Austronesian (Blust, this volume), where the distinction between inflection and derivation is especially hard to draw. On the other hand, the Bauer/Bakker and Geeraerts generalization about evaluative morphology has clear counter-examples as has been pointed out in recent works by Štekauer (2011) and Körtvélyessy (2012).

In this last section, we will see whether any of the material gathered in this volume gives us further insight into any of the proposed universals and cross-linguistic tendencies. Specifically, we have looked at 12 proposed universals/tendencies that are among the more easily testable ones, and have looked specifically for cases that seem to

contradict or problematize them. The ones we have looked at are summarized in (10), sometimes reworded a bit more succinctly than in the original:

- (10) a. If L has discontinuous affixes it always has either prefixes or suffixes or both (2a).
 b. If words can be derived from nouns, then they can be derived from non-nouns (3g).
 c. If L has nonconcatenative derivation then it also has concatenative derivation (3h).
 d. If L has augmentatives, it also has diminutives (4).
 e. All Ls have some sort of concatenative word-formation (6b).
 f. If L uses prefixation for word-formation, it also uses suffixation (7b).
 g. If L has derivational infixation, it also has derivational suffixation (7c).
 h. If L has derivational infixation, it also has derivational prefixation (7d).
 i. There is a strong tendency for Ls with suffixation to allow recursive suffixation (7e).
 j. There is a strong tendency for Ls to use reduplication as a word-formation process (8a).
 k. If L has derivational reduplication, it will usually have full reduplication (8b).
 l. If L has partial reduplication, it will also have full reduplication (8c).

Roughly half of these generalizations hold up according to the material presented in this volume. With regard to (10a), we note that discontinuous affixes seem to be quite rare to begin with in the languages of the world, and that those that show evidence of it do seem to have prefixation and suffixation as well (see, for example, Chapter 30 on Austronesian, Blust this volume). Looking at (10b), we have encountered no languages in which it is possible to use nouns as bases (10b), but not other categories of words. We find as well that most languages that allow suffixation are able to suffix recursively (10i). In Mandarin, affixation is marginal at best, so we do not find recursive suffixation, but (10i) is stated as a tendency, so the existence of a counter-example to it is neither surprising nor particularly problematic. And as for reduplication (10j–10l), the evidence amassed in this volume corroborates that offered in the *World Atlas of Linguistic Structures* (WALS) (see Inkelas, this volume) that reduplication can be found in a majority of the world's languages. We note that the tendency of western linguists to regard reduplication as somewhat exotic might stem from its relative paucity in Indo-European, the language family that linguists are most familiar with. Further, of the languages that evidence reduplication mentioned in this volume, it does turn out that all have full reduplication, although some do not have partial reduplication.

Interestingly, however, we find some evidence in the chapters gathered here that cast doubt on other items in (10). For example, (10c) suggests that we should not find languages in which derivation is effected solely by nonconcatenative means, and (10e) suggests that word-formation in all languages must involve some sort of concatenation.

The Nilo-Saharan language Dinka, however, might turn out to be an exception to both: according to Dimmendaal (this volume and p.c.), Dinka has some concatenative inflections (a perfective marker, subject enclitics) but derivation is by nonconcatenative means.

Universal (10f) addresses the so-called “suffixing preference” in languages (Hawkins and Gilligan 1988, Bauer this volume). The material in Hawkins and Gilligan suggests that there might be a minority of languages that are exclusively prefixing (ten in their combined samples 1988: 228), although their sample mixes inflectional and derivational categories and does not claim to look at the complete morphology of any language. Štekauer et al. (2012) provide Yoruba as a counter-example to this universal. However, several chapters in this volume give us a clearer picture. The chapter by Alves suggests that Mon-Khmer languages do not have derivational suffixes, and Vajda’s chapter on Yenisean indicates a strong prefixing preference in that family as well, finding only one productive suffix. In Chapter 34, Sun argues as well that Rgyalrong strongly prefers prefixing. Alves’s chapter also calls into question (10g), as the Mon Khmer family has both prefixation and infixation, but not suffixation. Blevins also cites the case of Pingding Mandarin, which has an infix <|> but no other affixes. In light of the existence of affixes in other dialects of Mandarin (Chung, this volume), however, this is a point that might warrant further exploration.

What can we conclude then? First, we must return to the difficulties of doing typological work on derivation. Without a more intimate understanding of derivation in the language families we have reported on in this volume, and a more in-depth study of particular languages, it is difficult to say anything about some of the universals or universal tendencies mentioned in Section 41.3. We are clearly just at the beginning of typological research on derivation in the languages of the world, but we at least have some idea of the direction such research should take.

REFERENCES

- Abbi, A. (1992). *Reduplication in South Asian Languages: An Areal, Typological and Historical Study*. New Delhi: Allied Publishers.
- Ackema, P. and Neeleman, A. (2005). Word formation in Optimality Theory. In P. Štekauer and R. Lieber (eds), *Handbook of Word Formation*. Dordrecht: Springer, 285–315.
- Adams, V. (1973). *An Introduction to Modern English Word-formation*. London: Longman.
- Adams, V. (2001). *Complex Words in English*. Harlow: Longman.
- Agnihotri, R. K. (1998). Mixed codes and their acceptability. In R. K. Agnihotri, A. L. Khanna, and I. Sachdev (eds), *Social Psychological Perspectives on Second Language Learning*. New Delhi: Sage Publications, 215–230.
- Ahland, C. (2010). Noun incorporation and predicate classifiers in Gumuz. *Journal of African Languages and Linguistics* 31(2): 159–203.
- Aikhenvald, A. Y. (1996). Areal diffusion in northwest Amazonia: the case of Tariana. *Anthropological Linguistics* 38: 73–116.
- Aikhenvald, A. Y. (2002). *Language Contact in Amazonia*. New York: Oxford University Press.
- Aikhenvald, A. Y. (2003). Mechanisms of change in areal diffusion: new morphology and language contact. *Journal of Linguistics* 39: 1–29.
- Aikhenvald, A. Y. (2006). Grammars in contact: a cross-linguistic perspective. In A. Y. Aikhenvald, and R. M. W. Dixon (eds), *Grammars in Contact: A Cross-Linguistic Typology*. Oxford: Oxford University Press, 1–66.
- Aikhenvald, A. Y. (2007). Typological distinctions in word formation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Vol III: Grammatical Categories and the Lexicon*. 2nd edn. Cambridge: Cambridge University Press, 1–65.
- Aikhenvald, A. Y. (2008). *The Manambu Language of East Sepik, Papua New Guinea*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. (2012). *The Languages of the Amazon*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. and R. M. W. Dixon (eds) (2001). *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. and R. M. W. Dixon (eds) (2006). *Grammars in Contact: A Cross-Linguistic Typology*. Oxford: Oxford University Press, 143–162.
- Aimard, P. (1975). *Les jeux de mots de l'enfant*. Villeurbanne: Simép Editions.
- Ajello, R. (1998). Armenian. In A. G. Ramat, and P. Ramat (eds), *The Indo-European Languages*. London: Routledge, 197–227.
- Akinlabi, A. (1994). Alignment constraints in ATR harmony. *Studies in the Linguistic Sciences* 24: 1–18.
- Akinlabi, A. (1996). Featural affixation. *Journal of Linguistics* 32: 239–289.
- Akinlabi, A. (2011). Featural affixes. In M. van Oostendorp, C. Ewen, E. Hume, and K. Rice (eds), *The Blackwell Companion to Phonology*. Oxford: Wiley Blackwell, 1945–1971.
- Alamin, S., Schneider-Blum, G. and Dimmendaal, G. J. (2012). Finding your way in Tima. In A. Mietzner and U. Claudi (eds), *Directionality in Grammar and Discourse: Case Studies from Africa*. Cologne: Rüdiger Köppe, 9–33.

- Alanne, E. (1972). Zur Rolle der syntaktischen Interferenz der verwandten und unverwandten Sprachen. *Neuphilologische Mitteilungen* 73: 568–574.
- Alaska Native Language Center. <<http://www.uaf.edu/anlc/languages/stats/index.xml>>.
- Alber, B. and Arndt-Lappe, S. (2012). Templatic and subtractive truncation. In J. Trommer (ed.), *The Morphology and Phonology of Exponence*. Oxford: Oxford University Press, 289–325.
- Alderete, J. D. (1999). Morphologically governed accent in optimality theory. Ph.D. thesis, University of Massachusetts, Amherst, MA.
- Alderete, J. D. (2001). Dominance effects as transderivational Anti-faithfulness. *Phonology* 18: 201–253.
- Alderete, J. D., Beckman, J., Benua, L., Gnanadesikan, A., McCarthy, J., and Urbanczyk, S. (1999). Reduplication with fixed segmentism. *Linguistic Inquiry* 30: 327–365.
- Alexiadou, A. (2001). *Functional Structure in Nominals: Nominalization, and Ergativity*. Amsterdam: Benjamins.
- Alexiadou, A. (2009). On the role of syntactic locality in morphological processes: the case of (Greek) derived nominals. In A. Giannakidou and M. Rathert (eds), *Quantification, Definiteness and Nominalization*. Oxford: Oxford University Press, 253–280.
- Alexiadou, A. (2010a). Nominalizations: a probe into the architecture of grammar. Part I: the nominalization puzzle. *Language and Linguistics Compass* 4: 496–511.
- Alexiadou, A. (2010b). Nominalizations: a probe into the architecture of grammar. Part II: the aspectual properties of nominalizations, and the lexicon vs. syntax debate. *Language and Linguistics Compass* 4: 512–523.
- Alexiadou, A. and Martin, F. (2012). Competing affixes as aspectual morphemes: the case of deadjectival nominalizations. Online Proceedings of the 8th Mediterranean Morphology Meeting.
- Alexiadou, A. and Schäfer, F. (2006). Instrument subjects are Agents or Causers. *Proceedings of West Coast Conference on Formal Linguistics* 25: 40–48.
- Alexiadou, A. and Schäfer, F. (2010). On the syntax of episodic vs. dispositional -er nominals. In A. Alexiadou and M. Rathert (eds), *The syntax of Nominalizations across Languages and Frameworks*. Berlin: Mouton de Gruyter, 9–38.
- Alexiadou, A., Anagnostopoulou, E., and Schäfer, F. (2006). The properties of anti-causatives crosslinguistically. In M. Frascarelli (ed.), *Phases of Interpretation*. Berlin: Mouton de Gruyter, 187–211.
- Alhoniemi, A. (1993). *Grammatik des Tscheremissischen (Mari)*. Hamburg: Buske.
- Alio, K., and Jungraithmayr, H. (1989). *Lexique Bidiya*. Frankfurt am Main: Klostermann.
- Allen, M. (1978). Morphological investigations. PhD dissertation. University of Connecticut, Storrs.
- Alpher, B. (1973). Son of ergative. PhD dissertation. Cornell University.
- Alpher, B. (1991). *Yir-Yoront Lexicon: Sketch and Dictionary of an Australian Language*. Berlin: Mouton de Gruyter.
- Alpher, B. (2004). Pama-Nyungan: phonological reconstruction and status as a phylogenetic group. In C. Bower and H. Koch (eds), *Australian Languages: Classification and the Comparative Method* (Current Issues in Linguistic Theory 249). Amsterdam: Benjamins, 93–126.
- Alsina, A. (1992). On the argument structure of causatives. *Linguistic Inquiry* 23: 517–555.
- Alsina, A. (1999). Where's the Mirror Principle? *The Linguistic Review* 16: 1–42.
- Alves, M. J. (2001a). Distributional properties of causative verbs in Mon-Khmer languages. *Mon-Khmer Studies* 31: 107–120.

- Alves, M. J. (2001b). What's so Chinese about Vietnamese? In G. Thurgood (ed.), *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society*. Berkeley: University of California Press, 221–241.
- Alves, M. J. (2004). *A Pacoh Grammar*. Canberra: Pacific Linguistics.
- Amador Rodríguez, L. A. (2009). *La derivación nominal en español. Nombres de agente, instrumento, lugar y acción*. Bern: Lang.
- Amberber, M., Baker, B., and Harvey, M. (eds) (2010). *Complex Predicates: Cross-linguistic Perspectives on Event Structure*. Cambridge: Cambridge University Press.
- Ambrazas, S. (1993). On the development of diminutives in the Baltic languages. *Linguistica Baltica* 2: 47–67.
- Ameka, F. (1999). The typology and semantics of complex nominal duplication in Ewe. *Anthropological Linguistics* 41: 75–106.
- Amery, R. (1993). Encoding new concepts in old languages: a case study of Kaurua, the language of the Adelaide Plains. *Australian Aboriginal Studies* 1: 33–47.
- Anbessa T. (2000). A grammar of Sidaama. Ph.D. dissertation. The Hebrew University of Jerusalem.
- Anceaux, L. C. (1965). *The Nimboran Language: Phonology and Morphology*. Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 44. 's-Gravenhage: Martinns-Nijhoff.
- Andersen, P. K. (1987). Zero-anaphora and related phenomena in Classical Tibetan. *Studies in Language* 11: 279–312.
- Andersen, T. (1988). Consonant alternation in the verbal morphology of Pāri. *Afrika und Übersee* 71: 63–113.
- Andersen, T. (1990). Vowel length in Western Nilotic languages. *Acta Linguistica Hafniensia*, 22: 5–26.
- Andersen, T. (1992–4). Morphological stratification in Dinka: on the alternations of vowel quality, vowel length and tone in the morphology of transitive verb roots in a monosyllabic language. *Studies in African Linguistics* 23(1): 1–63.
- Anderson, G. D. S. (2004). Advances in Proto-Munda reconstruction. *Mon-Khmer Studies* 34: 159–184.
- Anderson, G. D.S. (2007). *The Munda Verb: Typological Perspectives*. Berlin: Walter de Gruyter.
- Anderson, S. R. (1982). Where's morphology? *Linguistic Inquiry* 13: 571–612.
- Anderson, S. R. (1985). Typological distinctions in word formation. In T. Shopen (ed.), *Language Typology and Syntactic Description Vol III: Grammatical Categories and the Lexicon*. New York: Cambridge University Press, 4–56.
- Anderson, S. R. (1992). *A-morphous Morphology*. Cambridge: Cambridge University Press.
- Andersson, E. (1994). Swedish. In E. König and J. van der Auwera (eds), *The Germanic Languages*. London: Routledge, 271–312.
- Andrews, S. and Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language* 68: 279–296.
- Ansaldo, U., Don, J., and Pfau, R. (2010). Parts of speech: particulars, universals and theoretical constructs. In U. Ansaldo, J. Don, and R. Pfau (eds), *Parts of Speech: Empirical and Theoretical Advances*. Amsterdam: Benjamins, 1–4.
- Anshen, F., and M. Aronoff. (1999). Using dictionaries to study the mental lexicon. *Brain and Language* 68: 16–26.

- Anward, J. (2001). Parts of speech. In M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds), *Language Typology and Language Universals. Volume I*. Berlin: de Gruyter, 726–735.
- Appah, C. K. I. and Appiah Amfo, N. A. (2011). The morphopragmatics of the diminutive morpheme (-ba /-wa) in Akan. *Lexis 6: Diminutives and Augmentatives in the Languages of the World* 6: 59–84. Retrieved from <http://lexis.univ-lyon3.fr/IMG/pdf/Lexis_6.pdf>.
- Applegate, R. (1972). Ineseño Chumash grammar. Ph.D. dissertation. University of California.
- Applegate, R. (1976). Reduplication in Ineseño Chumash. In M. Langdon and S. Silver (eds), *Hokan Studies*. The Hague: Mouton, 271–283.
- Appleyard, D. (2001). The verb ‘to say’ as a verb ‘recycling device’ in Ethiopian languages. In A. Zaborski (ed.), *New Data and New Methods in Afroasiatic Linguistics: Robert Hetzron, in Memoriam*. Wiesbaden: Harrassowitz, 1–11.
- Apresjan, J. D. (1974). Regular polysemy. *Linguistics* 142: 5–32.
- Arad, M. (2003). Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 21: 737–778.
- Arad, M. (2005). *Roots and Patterns: Hebrew Morpho-syntax*. Dordrecht: Springer.
- Archangeli, D. (1983). The root CV-template as a property of the affix: evidence from Yawelmani. *Natural Language and Linguistic Theory* 1: 348–384.
- Archangeli, D. (1985). Yokuts harmony: evidence for coplanar representation in nonlinear phonology. *Linguistic Inquiry* 16: 335–272.
- Archangeli, D. (1991). Syllabification and prosodic templates in Yawelmani. *Natural Language and Linguistic Theory* 9: 231–283.
- Arenas, P. (2003). *Etnografía y Alimentación entre los Toba-Ñachilamoléek y Wichí-Lhuku’tas del Chaco Central (Argentina)*. Self-published. Buenos Aires.
- Arenas, P. and Scarpa, G. (2007). Edible wild plants of the Chorote Indians, Gran Chaco, Argentina. *Botanical Journal of the Linnean Society* 153: 73–85.
- Arnon, I. and Snider, N. (2010). More than words: frequency effects for multi-word phrases. *Journal of Memory and Language* 62(1): 67–82.
- Arnott, D. W. (1970). *The Nominal and Verbal system of Fula*. Oxford: Clarendon.
- Arntz, R., Picht, H. and Mayer, F. (2009). *Einführung in die Terminologearbeit*. 6th edn. Hildesheim: Olms.
- Aronoff, M. (1976). *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Aronoff, M. (1980). Contextuals. *Language* 56(4): 744–758.
- Aronoff, M. (1983). Potential words, actual words, productivity and frequency. *Proceedings of the 13th International Congress of Linguists* 163–171.
- Aronoff, M. (1988). Head operations and strata in reduplication: a linear treatment. *Yearbook of Morphology* 1: 1–15.
- Aronoff, M. (1994). *Morphology by Itself*. Cambridge, MA: MIT Press.
- Aronoff, M. and Anshen, F. (1998). Morphology and the lexicon: lexicalization and productivity. In A. Spencer and A. M. Zwicky (eds), *The Handbook of Morphology*. Oxford: Blackwell, 237–248.
- Aronoff, M., and S. Cho. (2001). The semantics of -ship suffixation. *Linguistic Inquiry* 32: 167–173.
- Aronoff, M. and Fudeman, K. (2005). *What Is Morphology?* Malden, MA: Blackwell.
- Aronoff, M. and Fudeman, K. (2011). *What Is Morphology?* 2nd edn. Oxford: Wiley-Blackwell.
- Aronoff, M. and Fuhrhop, N. (2002). Restricting suffix combinations in German and English: closing suffixes and the monosuffix constraint. *Natural Language and Linguistic Theory* 20: 451–490.

- Aronoff, M. and Schvaneveldt, R. (1978). Testing morphological productivity. *Annals of the New York Academy of Sciences* 318: 106–114.
- Artiagoitia, X., Ortiz de Urbina, J., and Hualde, J. I. (forthcoming). Basque. In P. Müller, I. Ohnheiser, S. Olsen and F. Rainer (eds), *Word-Formation: An International Handbook of the Languages of Europe*. Berlin: Mouton de Gruyter.
- Asher, N. (1993). *Abstract Objects in Discourse*. Dordrecht: Kluwer.
- Asher, R. E. and Kumari, T. C. (1997). *Malayalam*. London: Routledge.
- Audring, J. and F. Masini (forthcoming). *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press.
- Augst, G., Bauer, A., and Stein, A. (1977). *Grundwortschatz und Ideolekt. Empirische Untersuchung zur semantischen und lexicalischen Struktur des kindlichen Wortschatzes*. Tübingen: Max Niemeyer.
- Austin, P. K. (1981). *A Grammar of Diyari, South Australia* (Cambridge Studies in Linguistics 32). Cambridge: Cambridge University Press.
- Avrorin, V. (1959–61). *Grammatika nanajskogo jazyka* [A grammar of Nanaj]. Vols. 1–2. Moscow: Akademija Nauk.
- Avrorin, V. (1981). *Sintaksicheskie issledovanija po nanajskomu jazyku* [Studies on Nanai syntax]. Leningrad: Nauka.
- Avrorin, V. (2000). *Grammatika manchshurskogo pis'mennogo jazyka* [A grammar of Written Manchu]. Saint Petersburg: Nauka.
- Axelrod, M. (1993). *The Semantics of Time: Aspectual Categorization in Koyukon Athabaskan*. Lincoln: University of Nebraska Press.
- Axelrod, M. (1998). Lexis, grammar, and grammatical change: the Koyukon classifier prefixes. In M. Darnell (ed.), *Functionalism and Formalism in Linguistics*, vol. 2. Amsterdam: Benjamins, 39–58.
- Azeb Amha (2012). Omotic languages. In Z. Frajzyngier and E. Shay (eds), *The Afroasiatic Languages*. Cambridge: Cambridge University Press, 423–504.
- Baayen, R. H. (1989). A corpus-based approach to morphological productivity. Ph.D. dissertation. Free University of Amsterdam.
- Baayen, R. H. (1992). Quantitative aspects of morphological productivity. In G. Booij and J. van Marle (eds), *Yearbook of Morphology 1991*. Dordrecht: Kluwer, 109–149.
- Baayen, R. H. (1993). On frequency, transparency, and productivity. In G. Booij, J. van Marle (eds), *Yearbook of Morphology 1992*. Dordrecht: Kluwer, 181–208.
- Baayen, R. H. (1994a). Derivational productivity and text typology. *Journal of Quantitative Linguistics* 1: 16–34.
- Baayen, R. H. (1994b). Productivity in language production. *Language and Cognitive Processes* 9: 447–469.
- Baayen, R. H. (2003). Probabilistic approaches to morphology. In R. Bod, J. B. Hay, and S. Jannedy (eds), *Probability Theory in Linguistics*. Cambridge, MA: The MIT Press, 229–287.
- Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R. H. (2010a). Assessing the processing consequences of segment reduction in Dutch with naive discriminative learning. *Lingue e linguaggio* 9: 95–112.
- Baayen, R. H. (2010b). The directed compound graph of English: an exploration of lexical connectivity and its processing consequences. In S. Olsen (ed.), *New Impulses in Word-formation* (Linguistische Berichte Sonderheft 17). Hamburg: Buske, 383–402.

- Baayen, R. H. (2010c). A real experiment is a factorial experiment? *The Mental Lexicon* 5(1): 149–157.
- Baayen, R. H. (2012). Learning from the Bible: computational modelling of the costs of letter transpositions and letter exchanges in reading Classical Hebrew and Modern English. *Lingue e linguaggio* 11(2): 123–146.
- Baayen, R. H. and Lieber, R. (1991). Productivity and English derivation: a corpus based study. *Linguistics* 29: 801–843.
- Baayen, R. H. and Plag, I. (2007). Parsing is not weaknessless: suffix ordering revisited. Unpublished manuscript. Edmonton, Canada, and Siegen, Germany.
- Baayen, R. H. and Plag, I. (2009). Suffix ordering and morphological processing. *Language* 85(1): 109–152.
- Baayen, R. H. and Schreuder, R. (2000). Towards a psycholinguistic computational model for morphological parsing. *Philosophical Transactions of the Royal Society* (Series A: Mathematical, Physical and Engineering Sciences) 358: 1–13.
- Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- Baayen, R. H., Dijkstra, T., and Schreuder, R. (1997). Singulars and plurals in Dutch: evidence for a parallel dual route model. *Journal of Memory and Language* 53: 496–512.
- Baayen, R. H., Lieber, R., and Schreuder, R. (1997b). The morphological complexity of simplex nouns. *Linguistics* 35: 861–877.
- Baayen, R. H., Schreuder, R., De Jong, N. H., and Krott, A. (2002). Dutch inflection: the rules that prove the exception. In S. Nooteboom, F. Weerman, and F. Wijnen (eds), *Storage and Computation in the Language Faculty*. Dordrecht: Kluwer Academic Publishers, 61–92.
- Baayen, R. H., McQueen, J., Dijkstra, T., and Schreuder, R. (2003). Frequency effects in regular inflectional morphology: revisiting Dutch plurals. In R. H. Baayen, and R. Schreuder (eds), *Morphological Structure in Language Processing*. Berlin: Mouton de Gruyter, 355–390.
- Baayen, R. H., Wurm, L. H., and Aycok, J. (2007). Lexical dynamics for low-frequency complex words: a regression study across tasks and modalities. *The Mental Lexicon* 2: 419–463.
- Baayen, R. H., Levelt, W., Schreuder, R., and Ernestus, M. (2008). Paradigmatic structure in speech production. *Proceedings Chicago Linguistics Society* 43(1): 1–29.
- Baayen, R. H., Kuperman, V., and Bertram, R. (2010). Frequency effects in compound processing. In S. Scalise and I. Vogel (eds), *Cross-disciplinary Issues in Compounding*. Amsterdam: Benjamins.
- Baayen, R. H., Milin, P., Filipovic Durdjevic, D., Hendrix, P., and Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118(3): 438–481.
- Baayen, R. H., Hendrix, P., and Ramscar, M. (2012). Sidestepping the combinatorial explosion: towards a processing model based on discriminative learning. *Language and Speech* 56: 329–347.
- Baayen, R. H., Kuperman, V., Shaoul, C., and Milin, P., Kliegl, R., and Ramscar, M. (2013). Decomposition makes things worse: a discrimination learning approach to the time course of understanding compounds in reading. Manuscript, University of Tübingen and McMaster University.
- Badecker, W. and Caramazza, A. (1989). A lexical distinction between inflection and derivation. *Linguistic Inquiry* 20: 108–116.
- Baerman, M. (in press). *The Oxford Handbook of Inflection*. Oxford: Oxford University Press.
- Baeskow, H. (2010). His Lordship's -ship and the King of Golfdom: against a purely functional analysis of suffixhood. *Word Structure* 3(1): 1–30.

- Bakema, P. and Geeraerts, D. (2000). Diminution and augmentation. In G. Booij, C. Lehmann, and J. Mugdan (eds), *Morphologie/Morphology. Ein internationales Handbuch zur Flexion und Wortbildung/An International Handbook on Inflection and Word formation*. Vol. 1. Berlin: Mouton de Gruyter, 1045–1052.
- Baker, M. C. (1985a). Incorporation: a theory of grammatical function changing. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, MA.
- Baker, M. C. (1985b). The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373–415.
- Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Baker, M. C. (1996). *The Polysynthesis Parameter*. New York: Oxford University Press.
- Baker, M. C. (2003). *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.
- Baker, M. C. and N. Vinokurova. (2009). On agent nominalizations and why they are not like event nominalizations. *Language* 85: 517–556.
- Baker, M. C., Johnson, K., and Roberts, I. (1989). Passive arguments raised. *Linguistic Inquiry* 20: 219–251.
- Bakker, D. (2011). Language sampling. In J.-J. Song (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 100–127.
- Bakker, P. (2000). Convergence intertwining: an alternative way towards the genesis of mixed languages. In D. G. Gilbers, J. Nerbonne, and J. Schaeken (eds), *Languages in Contact* (Studies in Slavic and General Linguistics, 28). Amsterdam: Rodopi, 29–35.
- Balakaev, M. (ed.) (1962). *Sovremennyy kazaxskij jazyk. Fonetika i morfologija*. [Modern Kazakh. Phonetics and morphology.] Alma Ata: Akademija Nauk.
- Balling, L. and Baayen, R. H. (2008). Morphological effects in auditory word recognition: evidence from Danish. *Language and Cognitive Processes* 23: 1159–1190.
- Balling, L. and Baayen, R. H. (2012). Probability and surprisal in auditory comprehension of morpho-logically complex words. *Cognition* 125: 80–106.
- Balota, D., Cortese, M., and Pilotti, M. (1999). Visual lexical decision latencies for 2906 words. Available at: <http://www.artsci.wustl.edu/~dbalota/lexical_decision.html>.
- Balota, D. A., Pilotti, M., and Cortese, M. J. (2001). Subjective frequency estimates of 2,938 monosyllabic words. *Memory and Cognition* 29: 639–647.
- Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., and Yap, M. (2004). Visual word recognition for single-syllable words. *Journal of Experimental Psychology: General* 133: 283–316.
- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., and Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods* 39(3): 445–459.
- Bamgbose, A. (1966). *A Grammar of Yoruba*. Cambridge: Cambridge University Press.
- Banker, E. M. (1964a). Bahnar affixation. *Mon-Khmer Studies* 1: 99–117.
- Banker, E. M. (1964b). Bahnar reduplication. *Mon-Khmer Studies* 1: 119–134.
- Banksira, D. P. (2000). *Sound Mutations: The Morphophonology of Chaha*. Amsterdam: Benjamins.
- Bannard, C. and Matthews, D. (2008). Stored word sequences in language learning: the effect of familiarity on children's repetition of four-word combinations. *Psychological Science* 19: 241–248.
- Bareiro, V. (2006). *La etnia manjui lumnanas del Chaco*. Biblioteca Paraguaya de Antropología, vol. 56. Asunción: CEADUC.

- Barker, C. (1998). Episodic *-ee* in English: a thematic role constraint on new word formation. *Language* 74: 695–727.
- Barreteau, D. (1988). *Description du Mofu-Gudur*, vols. 1 and 2. Paris: ORSTOM.
- Bartens, R. (2000). *Permiläisten kielten rakenne ja kehitys*. [Structure and development of the Permic languages.] Mémoires de la Société Finno-Ougrienne 238. Helsinki: Suomalais-Ugrilainen Seura.
- Baskakov, N. A. (1962). *Vvedenie v izuchenie tjurkskix jazykov*. [Introduction to the study of the Turkic languages.] Moscow: Nauka.
- Bassène, A.-C. (2007). *Morphosyntaxe du Jóola Banjal*. Cologne: Rüdiger Köppe.
- Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. *Natural Language and Linguistic Theory* 12: 571–596.
- Bat-El, O. (2011). Semitic templates. In M. van Oostendorp, C. Ewen, E. Hume, and K. Rice (eds), *The Blackwell Companion to Phonology*. Oxford: Wiley Blackwell, 2586–2608.
- Bates, E. and Rankin, J. (1979). Morphological development in Italian: connotation and denotation. *Journal of Child Language* 6: 29–52.
- Bauer, A. H. (2007). Old English *-fæst*: a case of grammaticalisation. *Folia Linguistica Historica: Acta Societatis Linguisticae Europaeae* 28(1–2): 27–53.
- Bauer, C. (1982). Morphology and syntax of spoken Mon. Ph.D. dissertation. University of London. School of Oriental and African Studies, Department of the Languages and Cultures of Southeast Asia and the Islands.
- Bauer, C. (1989). The verb in spoken Mon. *Mon-Khmer Studies* 15: 87–110.
- Bauer, C. (1990). Reanalyzing reanalyses in Katuic and Bahnaric. *Mon-Khmer Studies* 16–17: 143–182.
- Bauer, L. (1983). *English Word-Formation*. Cambridge: Cambridge University Press.
- Bauer, L. (1988a). A descriptive gap in morphology. *Yearbook of Morphology* 1988. Dordrecht: Foris, 17–27.
- Bauer, L. (1988b). *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.
- Bauer, L. (1996). No phonetic iconicity in evaluative morphology. *Studia Linguistica* 50: 189–206.
- Bauer, L. (1997a). Evaluative morphology: in search of universals. *Studies in Language* 21(3): 533–575.
- Bauer, L. (1997b). Derivational Paradigms. In G. Booij and J. van Marle (eds), *Yearbook of Morphology* 1996. Dordrecht: Kluwer, 243–256.
- Bauer, L. (1998). Is there a class of neoclassical compounds, and if so is it productive? *Linguistics* 36(3): 403–422.
- Bauer, L. (2001). *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bauer, L. (2003). *Introducing Linguistic Morphology*. 2nd edn. Washington, DC: Georgetown University Press and Edinburgh: Edinburgh University Press.
- Bauer, L. (2004a). *A Glossary of Morphology*. Edinburgh: Edinburgh University Press.
- Bauer, L. (2004b). The function of word-formation and the inflection–derivation distinction. In H. Aertsen, M. Hannay, and R. Lyall (eds), *Words in their Places: A Festschrift for J. Lachlan Mackenzie*. Amsterdam: Vrije Universiteit, 283–292.
- Bauer, L. (2005a). The borderline between derivation and compounding. In W. U. Dressler, D. Kastovsky, O. E. Pfeiffer, and F. Rainer (eds), *Morphology and its Demarcations*. Amsterdam: Benjamins, 97–108.
- Bauer, L. (2005b). Conversion and the notion of lexical category. In L. Bauer and S. Valera (eds), *Approaches to Conversion/Zero-derivation*. Münster: Waxmann, 19–30.
- Bauer, L. (2010). An overview of morphological universals. *Word Structure* 3(2): 131–140.

- Bauer, L. and Huddleston, R. (2002). Lexical word-formation. In R. Huddleston and G. Pullum (eds), *The Cambridge Grammar of the English language*. Cambridge: Cambridge University Press, 1621–1721.
- Bauer, L., Lieber, R. and Plag, I. (2013). *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.
- Baxter, W. and L. Sagart (2012). Reconstructing the *s- prefix in Old Chinese. *Language and Linguistics* 13: 29–59.
- Beard, R. (1977). On the extent and nature of irregularity in the lexicon. *Lingua* 42: 305–341.
- Beard, R. (1981a). On the question of lexical regularity. *Journal of Linguistics* 17: 31–37.
- Beard, R. (1981b). *The Indo-European Lexicon*. Amsterdam: North-Holland.
- Beard, R. (1990). The nature and origins of derivational polysemy. *Lingua* 81: 101–140.
- Beard, R. (1995). *Lexeme-morpheme Base Morphology: A General Theory of Inflection and Word Formation*. Albany, NY: State University of New York Press.
- Beard, R. (1998). Derivation. In A. Spencer and A. M. Zwicky (eds), *The Handbook of Morphology*. Oxford: Blackwell, 44–65.
- Beavers, J. and Francez, I. (2012). Several problems for predicate decompositions. In Z. Anti, C. B. Chang, E. Cibelli, J. Hong, M. J. Houser, C. S. Sandy, M. Toosarvandani, and Y. Yao (eds), *Proceedings of Berkeley Linguistics Society* 32. Berkeley: University of California at Berkeley, 37–48.
- Beavers, J. and Koontz-Garboden, A. (in press). In defense of the reflexivization analysis of anticausativization. *Lingua*.
- Beck, D. (2002). The typology of parts of speech systems: the markedness of adjectives. In *Outstanding Dissertations in Linguistics*. New York: Routledge.
- Becker, J. A. (1994). ‘Sneak-shoes’, ‘swords’ and ‘nose-beards’: a case study of lexical innovation. *First Language* 14: 195–211.
- Becker, T. (1994). Back-formation, cross-formation, and ‘bracketing paradoxes’ in paradigmatic morphology. In G. Booij and J. van Marle (eds), *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 1–26.
- Beckwith, C. (1993). *The Tibetan Empire in Central Asia*. Princeton: Princeton University Press.
- Beckwith, C. and Walter, M. (2010). On the meaning of Old Tibetan rje-blon during the Tibetan Empire Period. *Journal Asiatique* 298(2): 535–548.
- Belletti, A. (2003). Agreement projections. In M. Baltin and C. Collins (eds), *Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, 483–510.
- Bender, M. L. (1996). *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom.
- Bendor-Samuel, J. (1960). Some problems of segmentation in the phonological analysis of Terena. *Word* 16: 348–155.
- Benedict, P.K. (1972). *Sino-Tibetan: A Conspectus*. Cambridge: Cambridge University Press.
- Benjamin, G. (1976). An outline of Temiar grammar. In P. N. Jenner, L. C. Thompson, and S. Starosta (eds), *Austroasiatic Studies Part II*, 129–188 (Oceanic Linguistics Special Publication No. 13). Honolulu: University of Hawaii Press.
- Benjamin, G. (2011). Temiar morphology (and Related Features): a view from the field. Fifth International Conference Austroasiatic Linguistics 5. Mahidol University, Salaya, Thailand.
- Bentin, S. and Frost, R. (2001). Linguistic theory and psychological reality: a reply to Boudelaa and Marslen-Wilson. *Cognition* 81(1): 113–118.
- Benveniste, É. (1948). *Noms d'agent et noms d'action en indo-européen*. Paris: Adrien-Maisonneuve.
- Benzing, J. (1955). *Lamutische Grammatik. Mit Bibliographie, Sprachproben und Glossar*. Wiesbaden: Franz Steiner Verlag GmbH.

- Berg, R. van den. (1996). The demise of focus and the spread of conjugated verbs in Sulawesi. In H. Steinhauer (ed.), *Papers in Austronesian Linguistics*. Canberra: Pacific Linguistics (PL A-84) 3: 89–114.
- Bergen, B. K. (2004). The psychological reality of phonaesthemes. *Language* 80: 290–311.
- Bergenholtz, H. and Mugdan, J. (1979). Ist *Liebe* primär? Über Ableitung und Wortarten. In Peter Braun (ed.), *Deutsche Gegenwartssprache*. München: Fink, 339–354.
- Bergsland, K. (1994). *Aleut Dictionary. Unungam Tunudgusii*. Fairbanks: Alaska Native Language Center.
- Bergsland, K. (1997). *Aleut Grammar. Unangam Tunuganaan Achixaasi*. Fairbanks: Alaska Native Language Center Research Paper Number 10, University of Alaska.
- Berko, J. (1958). The child's learning of English morphology. *Word* 14: 150–177.
- Berman, R. A. (1982). Verb-pattern alternation: the interface of morphology, syntax, and semantics in Hebrew child language. *Journal of Child Language* 9: 169–191.
- Berman, R. A. (1985). Acquisition of Hebrew. In D. I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition*, vol. 1. Hillsdale, NJ: Lawrence Erlbaum, 255–371.
- Berman, R. A. (1987). Productivity in the lexicon: new word formation in Modern Hebrew. *Folia Linguistica* XXI: 425–461.
- Berman, R. A. (2009). Children's acquisition of compound constructions. In R. Lieber and P. Stekauer (eds), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 298–322.
- Berman, Ruth A., and Sagi, Y. (1981). al darxey tecurat hamilim vexidushvan bagil haca'ir [Word-formation processes and lexical innovations of young children]. *Hebrew Computational Linguistics Bulletin* 18: 36–62.
- Bermúdez-Otero, R. (2012). The architecture of grammar and the division of labor in exponence. In J. Trommer (ed.), *The Morphology and Phonology of Exponence*. Oxford: Oxford University Press, 8–83.
- Bermúdez-Otero, R. (forthcoming). *Stratal Optimality Theory*. Oxford: Oxford University Press.
- Bertram, R., Baayen, R. H., and Schreuder, R. (2000a). Effects of family size for complex words. *Journal of Memory and Language* 42: 390–405.
- Bertram, R., Laine, M., and Virkkala, M. M. (2000c). The role of derivational morphology in vocabulary acquisition: get by with a little help from my morpheme friends. *Scandinavian Journal of Morphology* 41: 287–296.
- Bertram, R., Schreuder, R., and Baayen, R. H. (2000b). The balance of storage and computation in morphological processing: the role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: 419–511.
- Bessé, B. de (1997). Terminological Definitions [transl. J. C. Sager]. In S. E. Wright and G. Budin (eds), *Handbook of Terminology Management Volume 1: Basic Aspects of Terminology Management*. Amsterdam: Benjamins, 63–74.
- Beyer, S. (1992). *The Classical Tibetan language*. New York: State University of New York Press.
- Bhat, D. N. (1994). *The Adjectival Category: Criteria for Differentiation and Identification*. Amsterdam: Benjamins.
- Bick, A., Goelman, G., and Frost, R. (2008). Neural correlates of morphological processes in Hebrew. *Journal of Cognitive Neuroscience* 20(3): 406–420.
- Bickel, B. and Nichols, J. (2007). Inflectional morphology. In T. Shopen (ed.), *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press, 169–228.
- Biddle, N. (2012). *Indigenous Language Usage*. Canberra: Centre for Aboriginal Economic Policy Research, Australian National University.
- Bielec, D. (1998). *Polish: An Essential Grammar*. London: Routledge.

- Bien, H., Levelt, W., and Baayen, R. H. (2005). Frequency effects in compound production. *Proceedings of the National Academy of Sciences of the USA* 102: 17876–17881.
- Bien, H., Baayen, R. H., and Levelt, W. (2011). Frequency effects in the production of Dutch deverbal adjectives and inflected verbs. *Language and Cognitive Processes* 27: 683–715.
- Biese, Y. M. (1941). Origin and development of conversions in English. *Annales Academiae Scientiarum Fennicae B XLV*(2).
- Bird, H., Lambon R., Seidenberg, M., McClelland, J., and Patterson, K. (2003). Deficits in phonology and past-tense morphology: what's the connection? *Journal of Memory and Language* 48: 502–526.
- Bisang, W. (1996). Areal typology and grammaticalization: processes of grammaticalization based on nouns and verbs in East and Mainland South East Asian languages. *Studies in Language* 20(3): 519–597.
- Bisetto, A. and C. Melloni. (2007). Result nominals: a lexical semantic investigation. In online Proceedings of 5th Mediterranean Morphology Meeting.
- Bittner, M. (1987). On the semantics of the Greenlandic antipassive and related constructions. *International Journal of American Linguistics* 53: 194–231.
- Bittner, M. (2005). Future discourse in a tenseless language. *Journal of Semantics* 22(4): 339–387.
- Blackings, M., and N. Fabb. (2003). *A Grammar of Ma'di*. Berlin: Mouton de Gruyter.
- Blake, B. J. (1987). *Australian Aboriginal Grammar*. London: Croom Helm.
- Blake, B. J. (1991). *Australian Aboriginal Languages: A General Introduction*. 2nd edn. St Lucia: University of Queensland Press.
- Blake, B. J. and K. Burridge (eds) (2003). *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001* (Amsterdam Studies in the Theory and History of Linguistic Science, 237). Amsterdam: Benjamins.
- Blake, F. R. (1925). *A Grammar of the Tagalog Language*. New Haven, CT: American Oriental Society.
- Blažek, V. (2006). Current progress in Altaic etymology. *Linguistica Online* January 30.
- Blevins, J. (1996). Mokilese reduplication. *Linguistic Inquiry* 27: 523–530.
- Blevins, J. (1999). Untangling Leti infixation. *Oceanic Linguistics* 38(2): 383–403.
- Blevins, J. (2005). Yurok verb classes. *International Journal of American Linguistics* 71: 327–349.
- Blevins, J. (2012). Duality of patterning: absolute universal or statistical tendency? *Language and Cognition* 4(4): 275–296.
- Blevins, J. (forthcoming a). Periphrastic agreement in Yurok. In J. P. Blevins, F. Ackerman, and G. T. Stump (eds), *Paradigms and Periphrasis*. Stanford: CSLI.
- Blevins, J. (forthcoming b). Linguistic clues to Andamanese pre-history: understanding the North-South divide. In T. Güldemann, P. McConvell, and R. Rhodes (eds), *Hunter-gatherers and Linguistic History: A Global Perspective*. Cambridge: Cambridge University Press.
- Blevins, J. and Garrett, A. (1998). The origins of consonant-vowel metathesis. *Language* 74: 508–556.
- Blevins, J. P. (2001). Paradigmatic derivation. *Transactions of the Philological Society* 99: 211–222.
- Blevins, J. P. (2003). Stems and paradigms. *Language* 79: 737–767.
- Bloch, B. and Trager, G. L. (1942). *Outline of Linguistic Analysis*. Baltimore: Linguistic Society of America.
- Bloch-Trojnar, M. (2006). *Polyfunctionality in Morphology: A Study of Verbal Nouns in Modern Irish*. Lublin: Wydawnictwo KUL.
- Bloomfield, L. (1933). *Language*. New York: Henry Holt. [(1984). Chicago: University of Chicago Press.]

- Blust, R. (1979). Proto-Western Malayo-Polynesian vocatives. *Bijdragen tot de taal-, land- en volkenkunde* 135: 205–251.
- Blust, R. (1993). Kelabit-English vocabulary. *The Sarawak Museum Journal* 44.65 (NS): 141–226.
- Blust, R. (1998). Ca- reduplication and Proto-Austronesian grammar. *Oceanic Linguistics* 37: 29–64.
- Blust, R. (2003). *Thao Dictionary* (Language and Linguistics Monograph Series A5). Institute of Linguistics (Preparatory Office). Academia Sinica. Taipei: Academia Sinica.
- Blust, R. (2009). *The Austronesian Languages*. Canberra: Pacific Linguistics (PL 602).
- Blust, R. and S. Trussel. (2010). *Austronesian Comparative Dictionary*. web edition <www.trussel2.com/ACD revision 8/14/2012>.
- Bobaljik, J. D. (2012). *Universals in Comparative Morphology: Suppletion, Superlatives and the Structure of Words*. Cambridge, MA: MIT Press.
- Bochner, H. (1984). Inflection within derivation. *The Linguistic Review* 3(4): 411–442.
- Bodrogligeti, A. J. E. (2001). *A Grammar of Chagatay*. München: Lincom Europa.
- Bogoras, W. (1969). Chukchee. In F. Boas (ed.), *Handbook of American Indian Languages* (Bureau of Ethnology Bulletin 40, part 2). Charleston, SC: Nabu Press, 631–903.
- Bogoyavlenskij, D. N. (1973). The acquisition of Russian inflections. In C. A. Ferguson and D. I. Slobin (eds), *Studies of Child Language Development*. New York: Holt Rinehart and Winston, 284–292.
- Böhmerová, A. (2011). Suffixal diminutives and augmentatives in Slovak—a systemic view with some cross-linguistic considerations. *Lexis 6: Diminutives and Augmentatives in the Languages of the World* 6: 59–84. [Retrieved from <http://lexis.univ-lyon3.fr/IMG/pdf/Lexis_6_-_Bohmerova.pdf>.]
- Boldyrev, B. (1976). *Kategorija kosvennoj prinadležnosti v tungusomančzhurskix jazykax* [The category of alienable possession in Manchu-Tungus]. Moscow: Nauka.
- Boldyrev, B. (1987). *Slovoobrazovanie imen sushchestvitel'nyx v tungusomančzhurskix jazykax v sravnitel'no-istoricheskom osveshčenii* [Nominal word formation in Tungus-Manchu. A comparative aspect]. Novosibirsk: Nauka.
- Bolozky, S. (1994). On the formation of diminutives in Modern Hebrew morphology. *Hebrew Studies* 35: 47–63.
- Bolozky, S. (1999). *Measuring Productivity in Word Formation: The Case of Israeli Hebrew*. Leiden: Brill.
- Booij, G. (1986). Form and meaning in morphology: the case of Dutch agent nouns. *Linguistics* 24: 503–517.
- Booij, G. (1988). The relation between inheritance and argument linking: deverbal nouns in Dutch. In M. Everaert and A. Evers (eds), *Morphology and Modularity: In Honour of Henk Schultink*. Dordrecht: Foris, 57–73.
- Booij, G. (1996). Inherent versus contextual inflection and the split morphology hypothesis. In G. Booij and I. van Marle (eds), *Yearbook of Morphology*. Dordrecht: Springer, 1–16.
- Booij, G. E. (1997). Autonomous morphology and paradigmatic relations. In *Yearbook of Morphology 1996*. Dordrecht: Springer, 35–53.
- Booij, G. (2000). Inflection and derivation. In G. Booij, C. Lehmann and J. Mugdan (eds), *Morphologie/Morphology*. Berlin: Walter de Gruyter, 360–369.
- Booij, G. (2001). Lexical phonology: a review. In C. Kreidler (ed.), *Phonology: Critical Concepts*. London: Routledge, 460–490.
- Booij, G. (2002). *The Morphology of Dutch*. Oxford: Oxford University Press.
- Booij, G. (2006). Inflection and derivation. In K. Brown (ed.), *Encyclopedia of Language and Linguistics*. Oxford: Elsevier, 654–661.

- Booij, G. (2007). *The Grammar of Words*. 2nd edn. Oxford: Oxford University Press.
- Booij, G. E. (2008). Paradigmatic morphology. In B. Fradin (ed.), *La raison morphologique. Hommage à la mémoire de Danielle Corbin*. Amsterdam: Benjamins, 29–38.
- Booij, G. (2009). Compounding and construction morphology. In R. Lieber and P. Štekauer (eds), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 201–216.
- Booij, G. (2010). *Construction Morphology*. Oxford: Oxford University Press.
- Booij, G. and Lieber, R. (1993). On the simultaneity of morphological and prosodic structure. In S. Hargus and E. Kaisse (eds), *Phonology: Studies in Lexical Phonology* 4: 23–44.
- Booij, G. and Lieber, R. (2004). On the paradigmatic nature of affixal semantics in English and Dutch. *Linguistics* 42: 327–357.
- Booij, G., Lehmann, C. and Mugdan, J. (eds), *Morphologie/Morphology*. Berlin: Walter de Gruyter.
- Bopp, F. (1833). *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinischen, Litthauischen, Gothischen und Deutschen*. Berlin: Ferdinand Dümmler.
- Borer, H. (2005a). *In Name Only*. Oxford: Oxford University Press.
- Borer, H. (2005b). *The Normal Course of Events*. Oxford: Oxford University Press.
- Borer, H. (2013). *Taking Form*. Oxford: Oxford University Press.
- Boretzky, N., W. Enninger, and T. Stolz (eds) (1989). *Beiträge zum 5. Essener Kolloquium über 'Grammatikalisierung: Natürlichkeit und Systemökonomie'*. Vol. 1. (Essener Beiträge zur Sprachwandelforschung, 7). Bochum: Brockmeyer.
- Bornemann, E. and Risch, E. (1978). *Griechische Grammatik*. 2nd edn. Frankfurt am Main: Diesterweg.
- Borowiak, T. (2007). Mixed conjunct verbs and other manifestations of Hindi Englishization. *Investigationes Linguisticae* 15: 1–13. [http://www.staff.amu.edu.pl/~inveling/pdf/Borowiak_INVELING15.pdf] retrieved June 26, 2012.]
- Borsley, R. and Kornfilt, J. (2000). Mixed extended projections. In R. Borsley (ed.), *The Nature and Function of Syntactic Categories*. New York: Academic Press, 101–131.
- Botha, R. (1981). A base rule theory of Afrikaans synthetic compounding. In M. Moortgat, H. van der Hulst, and T. Hoekstra (eds), *The Scope of Lexical Rules*. Dordrecht: Foris, 1–77.
- Boudelaa, S. and Marslen-Wilson, W. D. (2001). Morphological units in the Arabic mental lexicon. *Cognition* 81(1): 65–92.
- Boudelaa, S. and Marslen-Wilson, W. D. (2004). Abstract morphemes and lexical representation: the cv-skeleton in Arabic. *Cognition* 92(3): 271–303.
- Boudelaa, S., Pulvermüller, F., Hauk, O., Shtyrov, Y., and Marslen-Wilson, W. (2009). Arabic morphology in the neural language system. *Journal of Cognitive Neuroscience* 22(5): 998–1010.
- Bourchier, L. (2008). Redupli-ca-cation. Poster presented at the 13th International Morphology Meeting, Vienna.
- Bowern, C. and Koch, H. (eds) (2004). *Australian Languages: Classification and the Comparative Method* (Current Issues in Linguistic Theory 249). Philadelphia: Benjamins.
- Boyd, J. (2007). Comparatively speaking: a psycholinguistic study of optionality in grammar. Ph.D. dissertation, City University of New York.
- Bozic, M., Marslen-Wilson, W. D., Stamatakis, E. A., Davis, M. H., and Tyler, L. K. (2007). Differentiating morphology, form, and meaning: neural correlates of morphological complexity. *Journal of Cognitive Neuroscience* 19(9): 1464–1475.
- Bram, B. (2011). Major total conversion in English: the question of directionality. Ph.D. dissertation. Victoria University of Wellington.

- Braunstein, J. and Vidal, A. (forthcoming). The Gran Chaco: convergence of languages and peoples. In R. Rhodes, T. Güldemann, and P. McConvell (eds), *The Languages of Hunter-gatherers. Historical and Global Perspectives*. Cambridge: Cambridge University Press.
- Bréal, M. (1868). *Les idées latentes du langage*. Paris: Hachette.
- Bréal, M. (1887). L'histoire des mots, *Revue des Deux Mondes*, LVIIe année—troisième période, tome quatre-vingt-deuxième, 187–212.
- Bredin, H. (1984). Sign and value in Saussure. *Philosophy* 59: 67–77.
- Brekke, H. H. (1978). Reflections on the conditions for coining, use, and understanding of nominal compounds. In W. U. Dressler and W. Meid (eds), *Proceedings of the XIIth International Congress of Linguists, Vienna 1977*. Innsbruck: Inbrucker Beitrage zur Sprachwissenschaft, 68–77.
- Bresnan, J., (ed.) (1982). *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.
- Bresnan, J. (1996). Lexicality and argument structure. Invited paper given at the Paris Syntax and Semantics Conference. October 12–14. <<http://www.stanford.edu/~bresnan/paris.pdf>>.
- Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.
- Bresnan, J. and Ford, M. (2010). Predicting syntax: processing dative constructions in American and Australian varieties of English. *Language* 86(1): 168–213.
- Bresnan, J. and Mchombo, S. A. (1995). The lexical integrity principle: evidence from Bantu. *Natural Language and Linguistic Theory* 13: 181–252.
- Bresnan, J. and Moshi, L. (1990). Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21: 147–185.
- Bresnan, J., Cueni, A., Nikitina, T., and Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Kramer, and J. Zwarts (eds), *Cognitive Foundations of Interpretation*. Amsterdam: Royal Netherlands Academy of Science, 69–94.
- Breu, W. (1996). Überlegungen zu einer Klassifizierung des grammatischen Wandels im Sprachkontakt (am Beispiel slavischer Kontaktfälle). *Sprachtypologie und Universalienforschung* 49(1): 21–38.
- Briggs, J., Johns, A., and Cook, C. (in progress). *Utkuhiksalingmiut Postbase Dictionary*. <<http://homes.chass.utoronto.ca/~inuit/UIDP/index.h>>.
- Brömser, B. (1985). On the derivation of English verbal compounds. In W. Kürschner and R. Vogt (eds), *Grammatik, Semantik, Textlinguistik*, Band 1. Tübingen: Niemeyer, 99–113.
- Broschart, J. (1997). Why Tongan does it differently: Categorical distinctions in a language without nouns and verbs. *Linguistic Typology* 1: 123–165.
- Brown, D. (2011). Morphological Typology. In J.-J. Song, (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 487–503.
- Brown, D. and Hippiusley, A. (2012). *Network Morphology*. Cambridge: Cambridge University Press.
- Brown, L., and Dryer, M. S. (n.d.). Diminutive as an inflectional category in Walman. Unpublished manuscript, Department of Linguistics, University at Buffalo, Buffalo, NY. Retrieved from <<http://linguistics.buffalo.edu/people/faculty/dryer/dryer/BrownDryerWalmanDimin.pdf>>.
- Brown, S. W. (2008). Polysemy in the mental lexicon. *Colorado Research in Linguistics* 21: 1–12.
- Brysbaert, M. and New, B. (2009). Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41(4): 977–990.
- Buba, B., and Owens, J. (2007). Glavda morphology. In A. S. Kaye (ed.), *Morphologies of Asia and Africa*, vol. 1. Winona Lake, IN: Eisenbrauns, 641–675.

- Burani, C. and Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes* 2: 217–227.
- Burani, C., Dovetto, M., Thornton, A. M., and Laudanna, A. (1997). Accessing and naming suffixed pseudo-words. In G. E. Booij, and J. Van Marle (eds), *Yearbook of Morphology 1996*. Dordrecht: Kluwer, 55–72.
- Burenhult, N. (2002). A grammar of Jahai. Ph.D. dissertation. Department of Linguistics and Phonetics, Lund University.
- Burnley, D. (1992). Word formation. In N. F. Blake (ed.), *The Cambridge History of the English Language* 2. 1066–1476. Cambridge: Cambridge University Press, 439–51.
- Butterworth, B. (1983). Lexical representation. In B. Butterworth (ed.), *Language production*, vol. 2. London: Academic Press, 257–294.
- Buzáßyová, K. (1974). *Sémantika slovenských deverbatív*. Bratislava: Veda.
- Bybee, J. L. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins.
- Bybee, J., Perkins, R., and Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Bye, P. and Svenonius, P. (2012). Non-concatenative morphology as epiphenomenon. In J. Trommer (ed.), *The Morphology and Phonology of Exponence*. Oxford: Oxford University Press, 42–495.
- Caballero, G. (2008). Choguita Rarámuri (Tarahumara) phonology and morphology. Ph.D. dissertation. University of California, Berkeley, CA.
- Caballero, G. (2010). Scope, phonology and morphology in an agglutinating language: Choguita Rarámuri (Tarahumara) variable suffix ordering. *Morphology* 20: 165–204.
- Caballero, G. (2011). Morphologically conditioned stress assignment in Choguita Rarámuri (Tarahumara). *Linguistics* 49(4): 749–790.
- Caballero, G. and Harris, A. C. (2012). A working typology of Multiple Exponence. In F. Kiefer et al. (eds), *Current Issues in Morphological Theory: (Ir)Regularity, Analogy and Frequency*. Selected papers from the 14th International Morphology Meeting, Budapest, May 13–16, 2010. Amsterdam: Benjamins.
- Campbell, L. (1987). Syntactic change in Pipil. *International Journal of American Linguistics* 53(3): 253–80.
- Campbell, L. (1997). *American Indian Languages: The Historical Linguistics of Native America*. Oxford: Oxford University Press.
- Campbell, L. and Langacker, R. (1978). Proto-Aztec Vowels, parts. I–III. *International Journal of American Linguistics* 44: 85–102, 197–210, 262–79.
- Caramazza, A., Miceli, G., Caterina, M., and Laudanna, A. (1985). Reading mechanisms and the organization of the lexicon: evidence from acquired dyslexia. *Cognitive Neuropsychology* 2: 81–114.
- Caramazza, A., Laudanna, A., and Romani, C. (1988). Lexical access and inflectional morphology. *Cognition* 28: 297–332.
- Carlson, R. (1994). *A Grammar of Supyire*. Berlin: Mouton de Gruyter.
- Carochi, H. (1645). Grammar of the Mexican language with an explanation of its adverbs. [Translated and edited with commentary by James Lockhart, UCLA Latin American Center.]
- Carol, J. (2009/10). Causación en chorote. *Amerindia* 33–34: 73–108.
- Carp, J. (2012). On the plurality of (methodological) worlds: estimating the analytic flexibility of fMRI experiments. *Frontiers in Neuroscience* 6: 1–13.
- Carrier, J. (1979). The interaction of phonological and morphological rules in Tagalog. Ph.D. dissertation, MIT.
- Carstairs, A. (1987). *Allomorphy in Inflexion*. London: Croom Helm.

- Carstairs, A. (1988). Some implications of phonologically conditioned suppletion. In G. Booij and J. van Marle (eds), *Yearbook of Morphology*. Dordrecht: Kluwer Academic Publishers, 67–94.
- Carstairs, A. (1990). Phonologically conditioned suppletion. In W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer, and J. R. Rennison (eds), *Contemporary Morphology*. New York: Mouton de Gruyter, 17–23.
- Carstairs-McCarthy, A. (1992). *Current Morphology*. New York: Routledge.
- Carstairs-McCarthy, A. (1998). Phonological constraints on morphological rules. In A. Spencer and A. Zwicky (eds), *The Handbook of Morphology*. Oxford: Blackwell, 144–148.
- Carstairs-McCarthy, A. (2005). Basic terminology. In P. Štekauer and R. Lieber (eds), *Handbook of Word-Formation*. Dordrecht: Springer, 5–24.
- Carstairs-McCarthy, A. (2006). Internal modification. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, vol. 5. Oxford: Elsevier, 752–755.
- Casali, R. F. (2008). ATR harmony in African languages. *Language and Linguistics Compass* 2(3): 496–549.
- Casaus, M. (in preparation). Quantitative ethnobotany and acculturation among the Rarámuri of Choguita, Chihuahua, México. Ph.D. dissertation, Cornell University.
- Castrén, M. A. (1858). *Versuch einer jensisei-ostjakischen und kottischen Sprachlehre*. St. Petersburg.
- Cataldi, L. (n.d.). *Ngardi Vocabulary and Notes*. Canberra: AIATSIS.
- Cense, A. A. (1979). *Makassaars-Nederlands woordenboek*. Koninklijk Instituut voor taal-, land- en volkenkunde. The Hague: Nijhoff.
- Census of India (2010). <<http://www.censusindia.gov.in/default.aspx>>. Retrieved June 23, 2012.
- Centineo, G. (1995). The distribution of *si* in Italian transitive/inchoative pairs. *Proceedings of Semantics and Linguistic Theory* 5: 54–71.
- Cerin, M. (1994). The pronominal system of Yaráalde. B.A. dissertation. University of Melbourne.
- Chang, B. S. (1971). The Tibetan causative phonology. *Bulletin of the Institute of History and Philology, Academia Sinica* 42: 623–765.
- Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chapman, D. and Skousen, R. (2005). Analogical modeling and morphological change: the case of the adjectival negative prefix in English. *English Language and Linguistics* 9(2): 333–357.
- Chappell, H. (2001). Language contact and areal diffusion in Sinitic languages. In A. Y. Aikhenvald and R. M. W. Dixon (eds), *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press, 328–357.
- Charekov, S. L. (1990). *Prilagatel'nye v mongol'skix jazykax v sravnenii s drugimi altajskimi* [Adjectives in Mongolic in comparison with other Altaic languages]. Leningrad: Nauka.
- Charles, T. (2012). Two deaf fishermen: a Klallam story told by Thomas Charles, with glossing and grammatical notes. Accessed August 1, 2012 at <www.ling.unt.edu/~montler/Klallam/Fishermen/>.
- Charney, J. O. (1993). *A Grammar of Comanche*. Lincoln, NE: University of Nebraska Press.
- Cheng, L. and Sybesma, R. (eds) (2003). *The Second Glot International State-of-the-Article Book: The Latest in Linguistics*. Berlin: Mouton de Gruyter.
- Chierchia, G. (2004). A semantics for unaccusatives and its syntactic consequences. In A. Alexiadou, E. Anagnostopoulou, and M. Everaert (eds), *The Unaccusativity Puzzle*. Oxford: Oxford University Press, 22–59.
- Childs, G. T. (1994). African ideophones. In L. Hinton et al. (eds), *Sound Symbolism*. Cambridge: Cambridge University Press, 247–279.
- Childs, G. T. (2011). *A Grammar of Mani*. Berlin: Walter de Gruyter.

- Chmura-Klekotowa, M. (1964). Rozwój rozumienia budowy wyrazów i umiejętności tworzenia wyrazów nowych (analogicznych) u dzieci w wieku przedszkolnym [The development of understanding of the internal structure of words and the capacity to create new (analogic) words in preschool children]. *Psychologia Wychowawcza* 1: 403–418.
- Chmura-Klekotowa, M. (1970). Odbicie tendencji słowotwórczych języka polskiego w neologizmach dzieci [Reflection of derivational trends of the Polish language in children's neologisms]. *Prace Filologiczne* 20: 153–159.
- Chmura-Klekotowa, M. (1971). Neologizmy słowotwórcze w mowie dzieci [Morphological neologisms in the speech of children]. *Prace Filologiczne* 21: 99–235.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1970). Remarks on nominalizations. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*. Waltham, MA: Ginn, 184–221.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1995). *Minimalism*. Cambridge, MA: MIT Press.
- Chomsky, N. (2008). On phases. In R. Freidin, C. Otero, and M.-L. Zubizarreta (eds), *Foundational Issues in Linguistic Theory*. Cambridge, MA: MIT Press, 133–166.
- Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Christianson, K., Johnson, R., and Rayner, K. (2005). Letter transpositions within and across morphemes. *Journal of Experimental Psychology: Learning Memory and Cognition* 31(6): 1327–1339.
- Chung, K. S. (2006). *Mandarin Compound Verbs*. Taipei: Crane Publishing.
- Chung, S. (2012). Are lexical categories universal? The view from Chamorro. *Theoretical Linguistics* 1–56.
- Chung, S. and Ladusaw, W. (2003). *Restriction and Saturation*. Cambridge, MA: MIT Press.
- Churchward, C. M. (1940). *Rotuman Grammar and Dictionary*. Sydney: Australian Medical Publishing Co. [Repr. AMS Press, New York, 1978.]
- CIA-The World Factbook (n.d.) <www.cia.gov/cia/publications/factbook/geos/kr.html#People>.
- Cincius, V. (1947). *Očerki grammatiki évenskogo (lamutskogo) jazyka* [A grammar of Even (Lamut)]. Leningrad: Učpedgiz.
- Cincius, V. (1949). *Sravnitel'naja fonetika tunguso-man'chzhurskix jazykov* [The comparative phonetics of the Tungus-Manchu languages]. Leningrad: Učpedgiz.
- Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-linguistic Perspective* (Oxford Studies in Comparative Syntax). Oxford: Oxford University Press.
- Ciszek, E. (2002). ME -lich(e)/-ly. *Studia Anglica Posnaniensia: An International Review of English Studies* 38: 105–129.
- Ciszek, E. (2004). On some French elements in Early Middle English word derivation. *Studia Anglica Posnaniensia: An International Review of English Studies* 40: 111–119.
- Ciszek, E. (2006). -Dōm in Medieval English. In N. Ritt et al. (eds), *Medieval English and Its Heritage: Structure, Meaning and Mechanisms of Change* (Studies in English Medieval Language and Literature: 16). Frankfurt: Peter Lang, 105–24.
- Ciszek, E. (2008). *Word Derivation in Early Middle English* (Studies in English Medieval Language and Literature: 23). Frankfurt: Peter Lang.
- Ciszek, E. (2012). The Middle English suffix -ish: reasons for decline in productivity <<http://www.thefreelibrary.com/The+Middle+English+suffix+-ish%3A+reasons+for+decline+in+productivity.-a0304842811>>, accessed February 14, 2013.

- Clackson, J. (2007). *Indo-European Linguistics: An Introduction*. Cambridge: Cambridge University Press.
- Clark, E. V. (1978). Awareness of language: Some evidence from what children say and do. In A. Sinclair, R. J. Jarvella, and W. J. M. Levelt (eds), *The Child's Conception of Language*. Berlin: SpringerVerlag, 17–43.
- Clark, E. V. (1981). Negative verbs in children's speech. In W. Klein and W. J. M. Levelt (eds), *Crossing the Boundaries in Linguistics: Studies Presented to Manfred Bierwisch*. Dordrecht: Reidel, 253–264.
- Clark, E. V. (1982). The young word-maker: A case study of innovation in the child's lexicon. In E. Wanner and L. R. Gleitman (eds), *Language Acquisition: The State of the Art*. Cambridge: Cambridge University Press, 390–425.
- Clark, E. V. (1984). Acquiring compounds. In *Proceedings of the First Eastern States Conference on Linguistics*. Columbus, OH: The Ohio State University, 181–190.
- Clark, E. V. (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clark, E. V. (2001). Emergent categories in first language acquisition. In M. Bowerman and S. C. Levinson (eds), *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press, 379–405.
- Clark, E. V. (2009). *First Language Acquisition*. Cambridge: Cambridge University Press.
- Clark, E. V. and Berman, R. A. (1984). Structure and use in the acquisition of word formation. *Language* 60: 542–590.
- Clark, E. V. and Clark, H. H. (1979). When nouns surface as verbs. *Language* 55: 767–811.
- Clark, E. V. and Cohen, S. R. (1984). Productivity and memory for newly formed words. *Journal of Child Language* 11: 611–625.
- Clark, E. V. and Hecht, B. F. (1982). Learning to coin agent and instrument nouns. *Cognition* 12: 1–24.
- Clark, E. V., Carpenter, K. L., and Deutsch, W. (1995). Reference states and reversals: undoing actions with verbs. *Journal of Child Language* 22: 633–662.
- Clark, E. V., Hecht, B. F., and Mulford, R. (1986). Coining complex compounds in English: affixes and word order in acquisition. *Linguistics* 24: 7–29.
- Clark, M. (n.d.). Notes on the causative prefix in some Mon-Khmer languages. University of Hawaii. ms.
- Clauson, G. (1959). The case for the Altaic theory examined. In H. Franke (ed.), *Akten des vierundzwanzigsten internationalen Orientalisten-Kongresses*. Wiesbaden: Deutsche Morgenländische Gesellschaft, in Kommission bei Franz Steiner Verlag.
- Clauson, G. (1968). A lexicostatistical appraisal of the Altaic theory. *Central Asiatic Journal* 13: 1–23.
- Coblin, W. S. (1976). Notes on Tibetan verb morphology. *T'oung Pao Archives* 62: 45–70.
- Cohen, P. D. (1966). Presyllables and reduplication in Jeh. *Mon-Khmer Studies* 2: 31–53.
- Cohn, A. (1989). Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7: 167–216.
- Cole, P., Beauvillain, C., and Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: a differential frequency effect. *Journal of Memory and Language* 28: 1–13.
- Coleman, D. (2012). *Mazin Grace*. St Lucia: University of Queensland Press.
- Collinder, B. (1960). *Comparative Grammar of the Uralic languages*. Stockholm: Almqvist and Wiksell.
- Comfort, J. and Jakobi, A. (2011). The verb 'to give' as a verbal extension in Uncunwee (Kordofan Nubian). In R. Kramer, H. Tröbs, and R. Kastenholz (eds), *Afrikanische Sprachen im Fokus*. Cologne: Rüdiger Köppe, 27–35.

- Compton, R. (2012). The syntax and semantics of modification in Inuktitut: adjectives and adverbs in a polysynthetic language. Ph.D. Thesis. University of Toronto.
- Compton, R. and Pittman, C. (2010). Word formation by phase in Inuit. *Lingua* 120: 2167–2192.
- Comrie, B. (1981). *The Languages of the Soviet Union*. Cambridge: Cambridge University Press.
- Comrie, B. and Thompson, S. A. (2007). Lexical nominalization. In T. Shopen (ed.), *Language Typology and Syntactic Description*. 2nd edn. Cambridge: Cambridge University Press, 334–381.
- Comrie, B., Golluscio L., González, L., and Vidal, A. (2010). El Chaco como área lingüística. In Z. Estrada and R. Arzapalo (eds.), *Estudios de lenguas amerindias 2. Lenguas indígenas*. Hermosillo: University of Sonora, 85–131.
- Conathan, L., and Good, J. (2000). Morphosyntactic reduplication in Chechen and Ingush. In A. Okrent and J. P. Boyle, (eds), *Proceedings of the Chicago Linguistic Society 36-2: The panels*. Chicago: Chicago Linguistic Society, 49–61.
- Condoravdi, C. and Kiparsky, P. (1998). Optimal order and scope. Paper presented at The Lexicon in Focus, Heinrich-Heine Universität Düsseldorf, August.
- Conrady, A. (1896). *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten*. Leipzig: O. Harrassowitz.
- Cook, C. and Johns, A. (2009). Determining the semantics of Inuktitut postbases. In M. A. Mahieu and N. Tersi (eds), *Variations on Polysynthesis: The Eskaleut Languages*. Amsterdam: Benjamins, 149–170.
- Cooper, J. S. (1966). Halång verb phrase. In D. D. Thomas (ed.), *Papers on Four Vietnamese Languages*. Auckland: Linguistic Society of New Zealand, 28–34.
- Cooreman, A. (1994). A functional typology of antipassives. In B. A. Fox and P. J. Hopper (eds), *Voice: Form and Function*. Amsterdam: Benjamins, 49–88.
- Corbett, G. (1987). The morphology–syntax interface: evidence from possessive adjectives in Slavonic. *Language* 63: 299–345.
- Corbett, G. (1999). Prototypical inflection: implications for typology. *Morphology* 1–22.
- Corbett, G. (2010). Canonical derivational morphology. *Word Structure* 3: 141–155.
- Corbett, G. G. (forthcoming). Canonical morphosyntactic features. In D. Brown, M. Chumakina, and G. G. Corbett (eds), *Canonical Morphology and Syntax*. Oxford: Oxford University Press.
- Corbin, D. (1987). *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer (2 vols).
- Corris, M. (2008). A grammar of Barupu: a language of Papua New Guinea. Ph.D. dissertation, University of Sydney.
- Corston-Oliver, S. (2002). Roviana. In J. Lynch, M. Ross, and T. Crowley (eds), *The Oceanic Languages*. Richmond: Curzon Press, 467–497.
- Cortina Borja M. and Valiñas Coalla, L. (1989). Some remarks on Uto-Aztecan classification. *International Journal of American Linguistics* 55: 214–239.
- Coseriu, E. (1952). *Sistema, norma y habla*. Montevideo: Universidad de la Republica, Facultad de Humanidades y Ciencias.
- Costa, S.R. (1976). Aprendizagem de alguns aspectos da morfologia portuguesa por crianças brasileiras [The learning of some aspects of Portuguese morphology by Brazilian children]. M.A. thesis. Universidade de Campinas, Brazil.
- Costello, N. A. (1966). Affixes in Katu. *Mon-Khmer Studies* 2: 63–86.
- Costello, N. A. (1998). Affixes in Katu of the Lao P.D.R. *Mon-Khmer Studies* 28: 31–42.
- Costello, N. A. (2001). Aspect and tense in Katu of the Lao P.D.R. *Mon-Khmer Studies* 31: 121–125.
- Coupe, A. R. (2007). *A Grammar of Mongsen Ao*. Berlin: Mouton de Gruyter.

- Cowie, C. (2012). Morphology. In A. Bergs and L. Brinton (eds), *English Historical Linguistics* 1(5): 604–621. [*Handbücher zur Sprach- und Kommunikationswissenschaft* 34.1].
- Craig, M. (1980). Aroma morphology. *Kivung* 12: 111–133.
- Crass, J. (2005). *Das K'abeena. Deskriptive Grammatik einer hochland-ostkuschitischen Sprache* (Kuschitische Sprachstudien 23). Cologne: Rüdiger Köppe.
- Crazzolar, J. P. (1938). *A Study of the Acoolli Language*. London: Oxford University Press for the International African Institute.
- Creider, C. A., and Creider, J. T. (2001). *A Dictionary of the Nandi Language*. Cologne: Rüdiger Köppe.
- Crick, F. H. C. (1989). The recent excitement about neural networks. *Nature* 337: 129–132.
- Croft, W. (1984). Semantic and pragmatic correlates to syntactic categories. In D. Testen, V. Mishra, and J. Drogo (eds), *Papers from the Parasession on Lexical Semantics*. Chicago, IL: Chicago Linguistic Society, 53–70.
- Croft, W. (1990). Possible verbs and the structure of events. In S. L. Tsohatzidis (ed.), *Meanings and Prototypes*. London: Routledge, 48–73.
- Croft, W. (1991). *Syntactic Categories and Grammatical Relations*. Chicago, IL: University of Chicago Press.
- Croft, W. (2000). Parts of speech as language universals and as language-particular categories. In P. M. Vogel and B. Comrie (eds), *Approaches to the Typology of Word-classes*. Berlin: Mouton de Gruyter, 65–102.
- Croft, W. (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Croft, W. and Poole, K. (2008). Inferring universals from grammatical variation: multidimensional scaling for typological analysis. *Theoretical Linguistics* 34: 1–37.
- Crowhurst, M. (1994). Foot extrametricality and template mapping in Cupeño. *Natural Language and Linguistic Theory* 12: 177–201.
- Crowley, T. (1978). *The Middle Clarence Dialects of Bandjalang*. Canberra: Australian Institute of Aboriginal Studies.
- Crowley, T. (1982). *The Paamese Language of Vanuatu*. Canberra: Pacific Linguistics (PL B-87).
- Crowley, T. (1993). Pre-1860 European contact in the Pacific and introduced cultural vocabulary. *Australian Journal of Linguistics* 13(2): 119–163.
- Crowley, T. (1998). *An Erromangan (Sye) grammar* (Oceanic Linguistics Special Publication 27). Honolulu: University of Hawaii Press.
- Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Crystal, D. (1967). Word classes in English. *Lingua* 17: 24–56.
- Cusihuaman, A. (1976). *Gramática Quechua: Cuzco-Collao*. Lima, Perú: Ministerio de Educación, Instituto de Estudios Peruanos.
- Cusic, D. (1981). Verbal plurality and aspect. Ph.D. dissertation, Stanford University.
- Cutler, A. (1980). Productivity in word formation. *Papers from the 16th regional meeting, Chicago Linguistic Society*. Chicago, IL: Department of Linguistics, University of Chicago, 45–51.
- Cutler, A., Hawkins, J. A. and Gilligan, G. (1985). The suffixing preference: a processing explanation. *Linguistics* 23: 723–758.
- Czaykowska-Higgins, E. and Kinkade, M. D. (1998). Salish languages and linguistics. In E. Czaykowska-Higgins and M. D. Kinkade (eds), *Salish Languages and Linguistics: Theoretical and Descriptive Perspectives*. Berlin: Mouton de Gruyter, 1–69.
- Dabrowska, E. (2006). Low level schemas or general rules? The role of diminutives in the acquisition of Polish case inflections. *Language Sciences* 28: 120–135.

- Dahl, Ö. (ed.) (2000). *Tense and Aspect in the Languages of Europe* (Empirical Approaches to Language Typology, Eurotyp, 20–6). Berlin: Mouton de Gruyter, 309–28.
- Daltas, P. (1987). Some patterns of variability in the use of diminutive and augmentative suffixes in spoken Modern Greek Koine (MGK). *Glossologia* 4: 63–88.
- Dalton-Puffer, C. (1994). Productive or not productive? The Romance element in Middle English derivation. In F. Fernández, M. Fuster, and J. J. Calvo (eds), *English Historical Linguistics 1992* (Amsterdam Studies in the Theory and History of Linguistic Science IV: Current Issues in Linguistic Theory: 113). Amsterdam: Benjamins, 247–60.
- Dalton-Puffer, C. (1996a). Suffixal derivation in Middle English: a corpus-based study. *Dissertation Abstracts International* 57: 3.
- Dalton-Puffer, C. (1996b). *The French Influence on Middle English Morphology. A Corpus-based Study of Derivation*. Berlin: Mouton de Gruyter.
- Dalton-Puffer, C. and Plag, I. (2000). Categorywise, some compound-type morphemes seem to be rather suffix-like: on the status of *-ful*, *-type* and *-wise* in Present-Day English. *Folia Linguistica* 34(3–4): 225–244.
- Dambacher, M. and Kliegl, R. (2007). Synchronizing timelines: relations between fixation durations and n400 amplitudes during sentence reading. *Brain Research* 1155: 147–162.
- Danchev, A. (1989). Language change typology and adjectival comparison in contact situations. *Folia Linguistica Historica* 9(2): 161–174.
- Danks, D. (2003). Equilibria of the Rescorla–Wagner model. *Journal of Mathematical Psychology* 47(2): 109–121.
- Darmesteter, A. (1887). *La vie des mots étudiée dans leurs significations*. Paris: Delagrave.
- Dasgupta, P. (2003). Bangla. In G. Cardona and D. Jain (eds), *The Indo-Aryan Languages*. London: Routledge, 351–90.
- Davidson, D. (1967). The Logical form of action sentences. In N. Resher (ed.), *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press, 81–95.
- Davis, K. (2003). *A Grammar of the Hoava Language, Western Solomons*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Davis, S. and Ueda, I. (2002). Mora augmentation processes in Japanese. *Journal of Japanese Linguistics* 18: 1–23.
- Davis, S. and Ueda, I. (2006). Prosodic vs. morphological mora augmentation. *Lexicon Forum* 2: 121–143.
- Davis, S. and Zawaydeh, B. A. (1999a). A descriptive analysis of hypocoristics in Colloquial Arabic. *Language and Linguistics* 3: 83–98.
- Davis, S. and Zawaydeh, B. A. (1999b). Hypocoristic formation in Ammani-Jordanian Arabic. In E. Benmamoun (ed.), *Perspectives on Arabic Linguistics XII*. Amsterdam: Benjamins, 113–139.
- Davis, S. and Zawaydeh, B. A. (2001). Arabic hypocoristics and the status of the consonantal root. *Linguistic Inquiry* 32: 512–520.
- Davliyeva, A. R. (2011). An investigation of Kazan Tatar morphology. M.A. thesis. University of San Diego. <http://sdsu-dspace.calstate.edu/bitstream/handle/10211.10/1063/Davliyeva_Albin.pdf?sequence=1>.
- Dayley, J. P. (1985). *Tzutujil Grammar*. Berkeley: University of California Press.
- Dayley, J. P. (1989). *Tumpisa (Panamint) Shoshone Grammar*. Berkeley: University of California Press.
- De Belder, M., Faust, N. and Lampitelli, N. (2010). On an inflectional and a derivational diminutive. Manuscript. Université Paris Diderot-Paris 7.

- de Groot, C. (2000). Minor word classes. In G. E. Booij, C. Lehmann, and J. Mugdan (eds), *Morphology: An International Handbook on Inflection and Word Formation*, vol. 1. Berlin: Walter de Gruyter, 820–831.
- De Jong, N. H., Schreuder, R., and Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes* 15: 329–365.
- De Jong, N. H., Schreuder, R., and Baayen, R. H. (2003). Morphological resonance in the mental lexicon. In R. H. Baayen and R. Schreuder (eds), *Morphological Structure in Language Processing*. Berlin: Mouton de Gruyter, 65–88.
- de Reuse, W. (1994). *Siberian Yupik Eskimo The Language and Its Contacts with Chukchi*. Salt Lake City: University of Utah Press.
- de Reuse, W. (2000). Polysynthesis as a typological feature: an attempt at a characterization from Eskimo and Athabascan perspectives. In M.-A. Mahieu and N. Tersis (eds), *Variations on Polysynthesis: The Eskaleut languages*. Amsterdam: Benjamins, 19–34.
- de Reuse, W. (2006). Polysynthetic language: Central Siberian Yupik. *Encyclopedia of Language and Linguistics*. 2nd edn. Amsterdam: Elsevier.
- De Schutter, G. (1994). Dutch. In E. König and J. van der Auwera (eds), *The Germanic Languages*. London: Routledge, 439–477.
- De Souza, I. C. (2010). A phonological description of ‘Pet Talk’ in Arara. M. A. thesis, University of North Dakota, Grand Forks.
- de Vaan, L., Schreuder, R. and Baayen, R.H. (2007). Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon* 2(1): 1–24.
- de Vooy, C. G. N. (1916). Iets over woordvorming en woordbetekenis in kindertaal. *De Nieuwe Taalgids* 10: 93–100, 128–141.
- Deacon, S. H. and Bryant, P. (2005). The strength of children’s knowledge of the role of root morphemes in the spelling of derived words. *Journal of Child Language* 32: 375–389.
- Deacon, S. H. and Bryant, P. (2006). Getting to the root: young writers’ sensitivity to the role of root morphemes in the spelling of inflected and derived words. *Journal of Child Language* 33: 401–417.
- Dedrick, J. M., and Casad, E. H. (eds) (1999). *Yaqui Language Structures*. Tucson: University of Arizona Press.
- DeFrancis, J. (1984). *The Chinese Language: Fact and Fantasy*. Honolulu: University of Hawaii Press. Taipei. [Crane Publishing reprint, 1990.]
- DeLancey, S. (1981). The category of direction in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 6(1): 83–101.
- Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93: 283–321.
- Delmas, C. (2000). (-)ABLE en anglais. In J. Pauchard (ed.), *La Modalité et les modaux en diachronie et en synchronie. Domaine anglais* (Recherches en Linguistique et Psychologie Cognitive: 14). Reims, France: PU de Reims, 151–173.
- Denison, D. (2001). Gradience and linguistic change. In L. J. Brinton (ed.), *Historical Linguistics 1999. Selected Papers from the 114th International Conference on Historical Linguistics, Vancouver, 9th–13th August 1999*. Amsterdam: Benjamins, 119–144.
- Denwood, P. (1986). The Tibetan noun-final -s. *Linguistics of the Tibeto-Burman Area* 9(1): 97–101.
- Denwood, P. (1999). *Tibetan*. Amsterdam: Benjamins.
- Deny, J. et al. (1959–64). *Philologiae Turcicae Fundamenta*. Wiesbaden: Harrassowitz.
- Derzhanski, I. A. (2005). On diminutive plurals and plural diminutives. *Morphology and Linguistic Typology*. On-line Proceedings of the Fourth Mediterranean Morphology Meeting

- (MMM4) Catania September 21–23, 2003, University of Bologna. Retrieved from <<http://morbo.lingue.unibo.it/mmm/>>.
- Descoeudres, A. (1922). *Le développement de l'enfant de deux à sept ans*. Neuchâtel: Delachaux et Niestlé.
- Detges, U. (forthcoming). The Romance adverbs in *-mente*: a case study in grammaticalization. In P. Müller, I. Ohnheiser, S. Olsen, and F. Rainer (eds), *Word-Formation: An International Handbook of the Languages of Europe*. Berlin: Mouton De Gruyter.
- Deutsch, A. and Meir, A. (2011). The role of the root morpheme in mediating word production in Hebrew. *Language and Cognitive Processes* 26(4–6): 716–744.
- Deutsch, A., Frost, R., and Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition* 24: 1238–1255.
- Devlin, J. T., Jamison, H. L., Matthews, P. M., and Gonnerman, L. M. (2004). Morphology and the internal structure of words. *Proceedings of the National Academy of Sciences of the United States of America* 101(41): 14984–14988.
- DGEEC (Dirección General de Estadística, Encuestas y Censos) (2003). *II Censo Nacional Indígena de Población y Viviendas 2002. Pueblos indígenas del Paraguay. Resultados finales*. Asunción: Dirección General de Estadística, Encuestas y Censos.
- Di Sciullo, A.-M. (2005). *Asymmetry in Morphology*. Cambridge, MA: MIT Press.
- Di Sciullo, A.-M. and Williams, E. (1987). *On the Definition of Word* (Linguistic Inquiry Monograph 14). Cambridge, MA: MIT Press.
- Diakonoff, I. M. (1988). *Afrasian Languages*. Moscow: Nauka.
- Diakonoff, I. M. and Kogan, L. E. (2007). Akkadian morphology. In A. S. Kaye (ed.), *Morphologies of Asia and Africa*. Winona Lake, IN: Eisenbrauns, 13–48.
- Diependaele, K., Sandra, D., and Grainger, J. (2005). Masked cross-modal morphological priming: unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes* 20: 75–114.
- Dietiker, S. R. (1983). *En bonne forme*. Lexington, MA: D. C. Heath.
- Dietz, K. (2004). Die altenglischen Präfixbildungen und ihre Charakteristik. *Anglia: Zeitschrift für Englische Philologie* 122(4): 561–613.
- Dietz, K. (2005). Die altenglischen Nominalpräfixe æ- und ö-, das Verbalpräfix a- und ihre althochdeutschen Entsprechungen. Ein Beitrag zur vergleichenden Wortbildung der altgermanischen Sprachen. *Sprachwissenschaft* 30(1): 1–47.
- Diffloth, G. (1976a). Expressives in Semai. *Oceanic Linguistics Special Publications* No. 13. *Austroasiatic Studies Part I*: 249–264.
- Diffloth, G. (1976b). Minor-syllable vocalism in Sino-Tibetan languages. *Oceanic Linguistics Special Publications*, No. 13. *Austroasiatic Studies Part I*: 229–247.
- Diffloth, G. (1979). Jah-Hut, an Austroasiatic language of Malaysia. *Southeast Asian Linguistic Studies* 2: 73–118.
- Diffloth, G. (2005). The contribution of linguistic palaeontology and Austroasiatic. In L. Sagart, R. Blench, and A. Sanchez-Mazas (eds), *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon, 77–80.
- Diffloth, G. and Zide, N. (1992). Austro-Asiatic Languages. In W. Bright et al. (eds), *International Encyclopedia of Linguistics*. New York: Oxford University Press, 137–142.
- Dimmendaal, G. J. (1982). The two morphological classes in Nilotic. In M. Bechhaus-Gerst and R. Vossen (eds), *Nilotic Studies*. Berlin: Dietrich Reimer, 269–309.
- Dimmendaal, G. (1983). *The Turkana Language*. Dordrecht: Foris Publications.

- Dimmendaal, G. J. (2000). Morphology. In B. Heine and D. Nurse (eds), *African Languages: An Introduction*. Cambridge: Cambridge University Press, 161–193.
- Dimmendaal, G. (2001). Language shift and morphological convergence in the Nilotic area. *SUGIA* 16–17: 83–124.
- Dimmendaal, G. (2005). Head marking, dependent marking and constituent order in the Nilotic area. In F. K. Erhard Voeltz (ed.), *Studies in African Linguistic Typology*. Amsterdam: Benjamins, 71–92.
- Dimmendaal, G. (2008). Language ecology and linguistic diversity on the African continent. *Language and Linguistics Compass* 2(5): 840–858.
- Dimmendaal, G. (2009). Datives in Nilotic. *Afrikanistik Online*. <urn:nbn:de:0009-10-23558>.
- Dimmendaal, G. (2010). Differential object marking in Nilo-Saharan. *Journal of African Languages and Linguistics* 31: 13–46.
- Dimmendaal, G. (2011). *Historical Linguistics and the Comparative Study of African Languages*. Amsterdam: Benjamins.
- Dimmendaal, G. (2012). Metrical structures: a neglected property of Nilotic (and other African language families). *Studies in Nilotic Linguistics* 5: 1–26.
- Dimmendaal, G. (2014). The grammar of knowledge in Tima. In A. Y. Aikhenvald and R. M. W. Dixon (eds), *The Grammar of Knowledge*. Oxford: Oxford University Press.
- Dimmendaal, G. (forthcoming a). Nilo-Saharan. In R. Vossen (ed.), *The Handbook of African Languages*. Oxford: Oxford University Press.
- Dinh, L. T. (2007). Reduplication in the M'ngong language. In M. Alves, P. Sidwell, and D. Gil (eds), *SEALS VIII: Papers from the 8th Meeting of the Southeast Asian Linguistics Society 1998*, 57–65.
- Dirven, R. (1999). Conversion as a conceptual metonymy of event schemata. In K.-U. Panther and G. Radden (eds), *Metonymy in Language and Thought*. Amsterdam: Benjamins, 275–287.
- Dixon, R. M. W. (ed.) (1976). *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies.
- Dixon, R. M. W. (1980). *The Languages of Australia*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. (1988). *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- Dixon, R. M. W. (2002). *Australian Languages: Their Origin and Development*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. (2010). *Basic Linguistic Theory, Vol. 1 Methodology*. Oxford: Oxford University Press.
- Doerfer, G. (1963). Bemerkungen zur Verwandtschaft der sog. altaische Sprachen. In G. Doerfer (ed.), *Türkische und mongolische Elemente im Neupersischen, Bd. I: Mongolische Elemente im Neupersischen*. Wiesbaden: Franz Steiner Verlag, 51–105.
- Doerfer, G. (1974). Ist das Japanische mit den altaischen Sprachen verwandt? *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 114(1): 103–142.
- Doerfer, G. (1978). Classification problems of Tungus. In G. Doerfer and M. Weiwers (eds), *Tungusica*, Vol. 1. Wiesbaden: Otto Harrassowitz, 1–26.
- Dokulil, M. (1962). *Tvoření slov v češtině I. Teorie odvozování slov*. Prague: Nakladatelství Československé akademie věd.
- Dokulil, M. (1968a). Zur Frage der Konversion und verwandter Wortbildungsvorgänge und -beziehungen. *Travaux Linguistiques de Prague* 3: 215–239.
- Dokulil, M. (1968b). Zur Frage der sog. Nullableitung. In H. E. Brekle and L. Lipka (eds), *Wortbildung, Syntax und Morphologie*. The Hague: Mouton, 55–64.
- Don, J. (1993). Morphological conversion. Ph.D. dissertation. Utrecht University, Utrecht.

- Don, J. (2003). A note on conversion in Dutch and German. In L. Cornips and P. Fikkert (eds), *Linguistics in the Netherlands 2003*, Amsterdam: Benjamins, 33–43.
- Don, J. (2004). Categories in the lexicon. *Linguistics* 42(5): 931–956.
- Don, J. (2005a). On conversion, relisting and zero-derivation. *SKASE Journal of Theoretical Linguistics* 2(2): 2–16.
- Don, J. (2005b). Roots, deverbal nouns and denominal verbs. In G. Booij, E. Guevara, A. Ralli, S. Sgroi and S. Scalise (eds), *Morphology and Linguistic Typology. On-line Proceedings of the Fourth Mediterranean Morphology Meeting (MMM4). Catania, September 21–23, 2003*, <<http://mmm.lingue.unibo.it/mmm-proc/MMM4/MMM4-Proceedings-full.pdf>> accessed May 31, 2012.
- Donaldson, B. (1994). Afrikaans. In E. König and J. van der Auwera (eds), *The Germanic Languages*. London: Routledge, 478–504.
- Donaldson, T. (1980). *Ngiyambaa: The Language of the Wangaaybuwan* (Cambridge Studies in Linguistics 29). Cambridge: Cambridge University Press.
- Donegan, P. (1993). Rhythm and vocalic drift in Munda and Mon-Khmer. *Linguistics of the Tibeto-Burman Area* 16(1): 1–42.
- Donegan, P. J. and Stampe, D. (1983). Rhythm and the holistic organization of language structure. In J. F. Richardson, M. Marks, and A. Chukerman (eds), *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*. Chicago: Chicago Linguistic Society, 337–353.
- Dorais, L.-J. (2010). *The Language of the Inuit: Syntax, Semantics and Society*. Montréal: McGill-Queens University Press.
- Doron, E. (2003). Agency and voice: the semantics of the Semitic templates. *Natural Language Semantics* 11: 1–67.
- Downer, G. B. (1959). Derivation by tone-change in Classical Chinese. *Bulletin of the School of Oriental and African Studies* 22: 258–290.
- Downing, L. J. (1997). Correspondence effects in Siswati reduplication. *Studies in the Linguistic Sciences* 25: 81–95.
- Downing, L. J. (1999a). Morphological constraints on Bantu reduplication. *Linguistic Analysis*, 29: 6–46.
- Downing, L. J. (1999b). Verbal reduplication in three Bantu languages. In R. Kager, H. van der Hulst, and W. Zonneveld, (eds), *The Prosody–Morphology Interface*. Cambridge: Cambridge University Press, 62–89.
- Downing, L. J. (2000). Morphological and prosodic constraints on Kinande verbal reduplication. *Phonology* 17: 1–38.
- Downing, L. (2006). *Canonical Forms in Prosodic Morphology*. Oxford: Oxford University Press.
- Dowty, D. (1979). *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel Publishing.
- Dowty, D. (1989). On the semantic content of the notion ‘thematic role’. In G. Chierchia, B. H. Partee, and R. Turner (eds), *Properties, types, and Meaning, Volume II: Semantic Issues*, Dordrecht: Kluwer, 69–130.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language* 67: 547–619.
- Dresher, B. E. (2009). The word in Tiberian Hebrew. In K. Hanson and S. Inkelas (eds), *The Nature of the Word: Essays in Honor of Paul Kiparsky*, Cambridge, MA: MIT Press, 95–111.
- Dresher, B. E. and Johns, A. (1996). The law of double consonants in Inuktitut. *Linguistica Atlantica* 17: 79–95.

- Dressler, W. U. (1968). *Studien zur verbalen Pluralität. Iterativum, Distributivum, Durativum, Intensivum in der allgemeinen Grammatik, im Lateinischen und Hethitischen*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Dressler, W. (1986). Explanation in natural morphology, illustrated with comparative and agent-noun formation. *Linguistics* 24: 519–548.
- Dressler, W. U. (1987). Word formation as part of natural morphology. In W. U. Dressler, W. Mayerthaler, O. Panagl, and W. U. Wurzel (eds), *Leitmotifs in Natural Morphology*. Amsterdam: Benjamins, 99–126.
- Dressler, W. (1988). Preferences vs. strict universals in morphology. In M. Hammond and M. Noonan, (eds), *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego, CA: Academic Press, 143–154.
- Dressler, W. U. (1989). Prototypical differences between inflection and derivation. *Zeitschrift für Sprachwissenschaft und Kommunikationsforschung* 42: 3–10.
- Dressler, W. (1994). Evidence from the first stages of morphology acquisition for linguistic theory: Extragrammatic morphology and diminutives. *Acta Linguistica Hafniensia: International Journal of Linguistics* 27: 91–108.
- Dressler, W. (2005). Word-formation in natural morphology. In P. Štekauer and R. Lieber (eds), *Handbook of Word-formation*. Dordrecht: Springer, 267–284.
- Dressler, W. and Karpf, A. (1995). The theoretical relevance of Pre- and Protomorphology in Language Acquisition. *Yearbook of Morphology*. Dordrecht: Kluwer.
- Dressler, W. and Ladányi, M. (2000). Productivity in word-formation: a morphological approach. *Acta Linguistica Hungarica* 47: 103–144.
- Dressler, W. and Merlini Barberesi, L. (1994). *Morphopragmatic—Diminutives and Intensifiers in Italian, German, and Other Languages*. Berlin: Mouton de Gruyter.
- Duanmu, S. and Zhang, H. (2010). Phonology and the Chinese Lexicon. CUNY Word Conference, CUNY, New York. a. Abstract: available at: <<http://www.cunyphonology-forum.net/WORDABSTRACTS/word-namedabstracts/duanmu-zhang.pdf>> accessed August 15, 2012; b. PowerPoint: available at: <<http://www-personal.umich.edu/~duanmu/CUNY10Word.pdf>> accessed August 15, 2012.
- Dubois, J. (1962). *Etude sur la derivation suffixale en français moderne et contemporain*. Paris: Librairie Larousse.
- Dudas, K. (1975). The phonology and morphology of modern Javanese. Ph.D. thesis, University of Illinois, Urbana-Champaign, IL.
- Dunabeitia, J. A., Perea, M., and Carreiras, M. (2007). Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition. *Cognition* 105(3): 691–703.
- Dunabeitia, J., Kinoshita, S., Carreiras, M., and Norris, D. (2011). Is morpho-orthographic decomposition purely orthographic? Evidence from masked priming in the same-different task. *Language and Cognitive Processes* 26(4–6): 509–529.
- Dunn, M. (1999). A grammar of Chukchi. PhD thesis. Australian National University.
- Durie, M. (1985). *A Grammar of Acehnese: On the Basis of a Dialect of North Aceh*. Dordrecht: Foris.
- Durie, M. and Ross, M. D. (eds) (1996). *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. New York: Oxford University Press.
- Dutkova-Cope, L. (2001). The language of Czech Moravians in Texas: do you know what *Párknu Káru u Hauza* means? *Southwest Journal of Linguistics* 20: 51–84.
- Dybo, A. V. (1996). *Semanticheskaja rekonstrukcija v altajskoj etimologii: somaticheskie terminy (plechevoj pojas)* [Semantic reconstruction in Altaic etymology: somatic terms]. Moskva: Shkola 'Jazyki russkoj kultury'.

- Dybo, A. V. (2000). Mir praaltajcev: Wörter und Sachen. In *Problemy izuchenija dalnego rodstva jazykov na rubezhe tretjego tysjacheletija*. Moskva: Rossijskij gosudarstvennyj gumanitarnyj univerzitet, 38–48.
- Dybo, A. V. and Starostin, G. S. (2008). In defence of the comparative method, or the end of the Vovin controversy. *Aspects of Comparative Linguistics* 3. Moscow: RSUH Publishers, 109–258.
- É. Kiss, K., Kiefer, F. and Siptár, P. (2003), *Új magyar nyelvtan* [New Hungarian Grammar], Budapest: Osiris, 189–284.
- Eckardt, R., von Heusinger, K., and Schwarze, C. (eds) (2003). *Words in Time: Diachronic Semantics from Different Points of View*. Berlin: Mouton de Gruyter.
- Egerod, S. (1965). Verb inflection in Atayal. *Lingua* 15: 251–282.
- Ehret, C. (2001). *A Historical-Comparative Reconstruction of Nilo-Saharan*. Cologne: Rüdiger Köppe.
- Eklund, A., Andersson, M., Josephson, C., Johannesson, M., and Knutsson, H. (2012). Does parametric fMRI analysis with SPM yield valid results? An empirical study of 1484 rest datasets. *NeuroImage* 61(3): 565–578.
- Elbers, L. (1988). New names from old words: related aspects of children's metaphors and word compounds. *Journal of Child Language* 15: 591–617.
- El'konin, D. B. (1973). General course of development in the child of the grammatical structure of the Russian language (according to A. N. Gvozdev). In C. A. Ferguson and D. I. Slobin (eds), *Studies of Child Language Development*. New York: Holt, Rinehart and Winston, 565–583.
- Elman, J. L. (2009). On the meaning of words and dinosaur bones: Lexical knowledge without a lexicon. *Cognitive Science* 33(4): 547–582.
- Embick, D. (2004). On the structure of resultative participles in English. *Linguistic Inquiry* 35: 355–392.
- Embick, D. (2009). Roots, states, stative passives. Handout of talk presented at Roots, Stuttgart, June 11.
- Embick, D. (2010). *Localism versus Globalism in Morphology and Phonology* (Linguistic Inquiry Monographs 60). Cambridge: MIT Press.
- Embick, D. and Marantz, A. (2008). Architecture and blocking. *Linguistic Inquiry* 39(1): 1–53.
- Emeneau, M. (1955). *Kolami: A Dravidian Language*. Berkeley: University of California Press.
- Emmerick, R. E. (1983). Some lexical items from the Siddhasāra. In E. Steinkellner and H. Tauscher (eds), *Contributions on Tibetan Language, History and Culture*. (Wiener Studien zur Tibetologie und Buddhismuskunde 10). Vienna: Arbeitskreis für Tibetische und Buddhistische Studien, Universität Wien, 61–68.
- Enfield, N. J. (2004). Adjectives in Lao. In R. M. W. Dixon and A. Y. Aikhenvald (eds), *Adjective Classes: A Cross-linguistic Typology*, Oxford: Oxford University Press, 323–347.
- Enfield, N. J. and Diffloth, G. (2009). Phonology and sketch grammar of Kri, a Vietic language of Laos. *Cahiers de Linguistique—Asie Orientale (CLAO)* 38(1): 3–69.
- Engelberg, M. (2002). Finnish: the communication of gender in Finnish. In M. Hellinger and H. Bußmann (eds), *Gender Across Languages: The Linguistic Representation of Women and Men*. Amsterdam: Benjamins, 109–132.
- Engels, J. (2003). Polysemie. In G. Ueding (ed.), *Historisches Wörterbuch der Rhetorik*. Vol. 6. Tübingen: Niemeyer, 1530–1537.
- Epps, P. (2006). The Vaupés melting pot: Tucanoan influence on Hup. In A. Y. Aikhenvald and R. M. W. Dixon (eds), *Grammars in Contact: A Cross-Linguistic Typology*. Oxford: Oxford University Press, 267–89.

- Erben, J. (1964). Deutsche Wortbildung in synchronischer und diachronischer Sicht. *Wirkendes Wort: Deutsche Sprache in Forschung und Lehre* 14: 83–93.
- Erben, J. (1972). Der sinnesarme Tristan: Zur Wortbildung des Adjektivs bei Gottfried von Straßburg. In *Paul und Braunes Beiträge zur Geschichte der deutschen Sprache und Literatur* 94 (T), Sonderheft, 182–91.
- Erben, J. (1977). Sprachgeschichte als Systemgeschichte. In Institut für Deutsche Sprache Mannheim (ed.), *Sprachwandel und Sprachgeschichtsschreibung im Deutschen Sprache der Gegenwart*, 41. Düsseldorf: Schwann.
- Erben, J. (1983). *Einführung in die deutsche Wortbildungslehre*. 2nd rev. edn, Berlin: Schmidt.
- Erelt, M. and Metslang, H. (2006). Estonian clause patterns—from Finno-Ugric to Standard Average European. *Linguistica Uralica* 4: 254–66.
- Ernestus, M. (2000). *Voice Assimilation and Segment Reduction in Casual Dutch. A Corpus-based Study of the Phonology–Phonetics Interface*. Utrecht: LOT.
- Ernestus, M., Baayen, R. H., and Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language* 81: 162–173.
- Ernst, T. (2002). *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.
- Ethnologue, Languages of China. Available at: <http://www.ethnologue.com/show_country.asp?name=cn> accessed August 15, 2012
- Ettinger, S. (1974). *Diminutiv- und Augmentativbildung: Regeln und Restriktionen*. Tübingen: Tübinger Beiträge zur Linguistik.
- Ettinger, S. (1980). *Form und Funktion in der Wortbildung: die Diminutiv- und Augmentativmodifikation im Lateinischen, Deutschen und Romanischen (Portugiesisch, Spanisch, Italienisch und Rumänisch)*. Tübingen: Gunter Narr Verlag.
- Evans, N. (1988). Arguments for Pama-Nyungan as a genetic subgroup, with particular reference to initial laminalization. In N. Evans and S. Johnson (eds), *Aboriginal Linguistics* 1 (1); Armidale: Department of Linguistics, University of New England, 91–110.
- Evans, N. (1992). Macassan loans in top end languages. *Australian Journal of Linguistics* 12: 45–91.
- Evans, N. (2003a). Context, culture and structuration in the languages of Australia. *Annual Review of Anthropology* 2003 32: 13–40.
- Evans, N. (2003b). An interesting couple: the semantic development of dyad morphemes, *Arbeitspapier Nr 47 (Neue Folge)*. Köln: Institut für Sprachwissenschaft Universität zu Köln.
- Evans, N. and Osada, T. (2005). Mundari: The myth of a language without word classes. *Linguistic Typology* 9: 351–390.
- Evsev'ev, M. E. (1934, 1963), *Osnovy mordovskoj grammatiki*. Izbrannye trudy. Saransk: Mordovskoe knižnoe izdatel'stvo.
- Fabb, N. (1984). *Syntactic affixation*. Ph.D. dissertation, MIT.
- Fabb, N. (1988). Suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory* 6(4): 527–539.
- Fabre, A. (2012 [2005]). Los pueblos del Gran Chaco y sus lenguas, segunda parte: Los Mataguayo [online updated version of the paper published in *Suplemento Antropológico* 40(2) (2005): 313–435. Asunción, Paraguay].
- Fábregas, A. (2007). The internal syntactic structure of relational adjectives. *Probus*, 17: 1–36.
- Fábregas, A. and Scalise, S. (2012), *Morphology: From Data to Theories*. Edinburgh: Edinburgh University Press.
- Fabricius, A. H. (1998). *A Comparative Survey of Reduplication in Australian Languages*. Munich: Lincom Europa.

- Fanego, T. (2006). The role of language standardization in the loss of hybrid gerunds in Modern English. In L. E. Breivik, S. Halverson, and K. Haugland (eds), *'These things write I vnto thee': Essays in honour of Bjørg Bækken*. Oslo: Novus, 93–110.
- Farrell, P. (2001). Functional shift as category underspecification. *English Language and Linguistics* 5(1): 109–130.
- Fedjuneva, G. v. (2000). *Komi émakyv artmöm* [Noun formation in Komi.]. Syktyvkar: Rossijasa Naukajas Akadmija [The Russian Academy of Sciences].
- Fedjuneva, G. V. and Punegova, G. V. (2002). *Komi kyvjyn tor'jakyv'jas artmöm: učebnoj posobie* [Word formation in Komi: a (university) textbook]. Syktyvkar: Syktyvkar-skij gosudartsvennyj universitet.
- Feldman, L. B. (2000). Are morphological effects distinguishable from the effects of shared meaning and shared form? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(6): 1431–1444.
- Feldman, L. and Pastizzo, M. (2003). Morphological facilitation: The role of semantic transparency and family size. In R. H. Baayen and R. Schreuder (eds), *Morphological Structure in Language Processing*. Berlin: Mouton de Gruyter, 233–258.
- Feldman, L. B., O'Connor, P. A., and Moscoso del Prado Martin, F. (2009). Early morphological processing is morpho-semantic and not simply morpho-orthographic: evidence from the masked priming paradigm. *Psychonomic Bulletin and Review* 16(4): 684–691.
- Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist* 66: 103–114.
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., and Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods* 42(2): 488–496.
- Figueira, R. A. (1984). On the development of the expression of causativity: a syntactic hypothesis. *Journal of Child Language* 11: 109–127.
- Filbeck, D. (1996). Couplet and duplication in Mal. *Mon-Khmer Studies* 26: 91–106.
- Finkenstaedt, T., Leisi, E., and Wolff, D. (1970). *A Chronological English dictionary: Listing 80000 Words in Order of their Earliest Known Occurrence*. Heidelberg: Winter.
- Finley, S. (2009). Morphemic harmony as featural correspondence. *Lingua* 119: 478–501.
- Fitch, W. T. (2010). *The Evolution of Language*. New York: Cambridge University Press.
- Fitzgerald, C. M. (2000). Vowel hiatus and faithfulness in Tohono O'odham reduplication. *Linguistic Inquiry* 31: 713–722.
- Fitzgerald, C. M. (2001). The morpheme-to-stress principle in Tohono O'odham. *Linguistics* 39: 941–972.
- Fleischer, W. (1982). *Wortbildung der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Fleischer, W. (2000). Die Klassifikation von Wortbildungsprozessen. In G. Booij, C. Lehmann, and J. Mugdan (eds), *Morphologie/Morphology*. Berlin: De Gruyter, 886–897.
- Fleischer, W. and Barz, I. (1995). *Wortbildung der deutschen Gegenwartssprache*. 2nd rev. edn. Tübingen: Niemeyer.
- Fleischer, W. and Barz, I. (2007). *Wortbildung der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Fleischer, W. and Barz, I. (2012). *Wortbildung der deutschen Gegenwartssprache*. 4th rev. edn. Berlin: Walter de Gruyter.
- Fleischman, S. (1977). *Cultural and Linguistic Factors in Word Formation: An Integrated Approach to the Development of the Suffix -age*. Berkeley: University of California Press.
- Flores Farfán, J. A. (2004). Notes on Nahuatl typological change. *Sprachtypologie und Universalienforschung* 57: 85–97.

- Fodor, J. (1970). Three reasons for not deriving 'kill' from 'cause to die'. *Linguistic Inquiry* 1: 429–438.
- Ford, M. A., Davis, M. H., and Marslen-Wilson, W. D. (2010). Derivational morphology and base morpheme frequency. *Journal of Memory and Language* 63(1): 117–130.
- Forster, K. (1999). The microgenesis of priming effects in lexical access. *Brain and Language* 68(1–2): 5–15.
- Forster, K. (2000). The potential for experimenter bias effects in word recognition experiments. *Memory and Cognition* 28: 1109–1115.
- Fortescue, M. (1980). Affix ordering in West Greenlandic derivational processes. *International Journal of American Linguistics* 46(4): 259–278.
- Fortescue, M. (1983). *A Comparative Manual of Affixes for the Inuit dialects of Greenland, Canada and Alaska*. Copenhagen: Meddelelser om Grønland.
- Fortescue, M. (1984). *West Greenlandic* (Croom Helm Descriptive Grammars). London: Croom Helm.
- Fortescue, M. (1992). Morphophonemic complexity and typological stability in a polysynthetic language family. *International Journal of American Linguistics* 58(2): 242–248.
- Fortescue, M. (1995). The historical source and typological position of ergativity in Eskimo languages. *Études/Inuit/Studies* 19(2): 61–75.
- Fortescue, M., Jacobson, S. and Kaplan, L. (2010). *Comparative Eskimo Dictionary: With Aleut Cognates*. 2nd edn, Fairbanks: Alaska Native Language Center Research Paper Number 9, University of Alaska.
- Fowler, C. S. (1983). Lexical Clues to Uto-Aztecan Prehistory. *International Journal of American Linguistics* 49: 224–257.
- Fradin, B. (2009). IE, Romance: French. In R. Lieber and P. Štekauer (eds), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 417–435.
- Frajzyngier, Z. (1981). Some rules concerning vowels in Chadic. *Bulletin of the School of Oriental and African Studies* 44(2): 334–348.
- Frajzyngier, Z. (1989). *A Grammar of Pero*. Berlin: Reimer.
- Frajzyngier, Z. (1993). *A Grammar of Mupun*. Berlin: Reimer.
- Frajzyngier, Z. (2001). *A Grammar of Lele* (Stanford Monographs in African Linguistics). Stanford: CSLI.
- Frajzyngier, Z. (2008). *A Grammar of Gidar*. Frankfurt: Peter Lang.
- Frajzyngier, Z. (2012). *A Grammar of Wandala*. Berlin: Mouton de Gruyter.
- Frajzyngier, Z. and Koops, R. (1989). Double epenthesis and N-class in Chadic. In Z. Frajzyngier (ed.), *Current Progress in Chadic Linguistics*. Amsterdam: Benjamins, 233–250.
- Frajzyngier, Z., with Shay, E. (2002). *A Grammar of Hdi*. Berlin: Mouton de Gruyter.
- Frajzyngier, Z. and Shay, E. (2012). Chadic. In Z. Frajzyngier and E. Shay (eds), *The Afroasiatic Languages*. Cambridge: Cambridge University Press, 236–341.
- Frajzyngier, Z. and Johnston, E. with Edwards, A. (2005). *A Grammar of Mina*. Berlin: Mouton de Gruyter.
- Francis, B., Robson, J., and Read, B. (2001). An analysis of undergraduate writing styles in the context of gender and achievement. *Studies in Higher Education* 26(3): 313–326.
- Francis, G. (2013). Publication bias in 'Red, Rank, and Romance in Women Viewing Men' by Elliot et al. (2010). *Journal of Experimental Psychology: General* 142(1): 292–296.
- Fraser-Knowles, J. and Oakes, M. (1996). A new Bundjalung language: Baryulgil Square Talk. In M. Sharpe (ed.), *An Introduction to the Yugambeh-Bundjalung Language and its Dialects* (Armidale: The Author), 133–55.

- Frauenfelder, U. H. and Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G. E. Booij and J. v. Marle (eds), *Yearbook of Morphology 1991*. Dordrecht: Kluwer, 165–183.
- Fromkin, V. (ed.) (2000). *Linguistics*. Oxford: Blackwell.
- Frost, R., Forster, K. I., and Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23: 829–856.
- Frost, R., Deutsch, A., and Forster, K. (2000a). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(3): 751–765.
- Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., and Marslen-Wilson, W. D. (2000b). Morphological priming: dissociation of phonological, semantic and morphological factors. *Memory and Cognition* 28(8): 1277–1288.
- Frost, R., Kugler, T., Deutsch, A., and Forster, K. (2005). Orthographic structure versus morpho-logical structure: principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31(6): 1293–1326.
- Fuller, E. (1977). Chru phonemes. *Pacific Linguistics A*, 48: 77–86.
- Furdík, J. (2004). *Slovenská slovotvorba*. Prešov: Náuka.
- Gaeta, L. (2005). Combinazioni di suffissi in Italiano. In M. Grossmann and A. M. Thornton (eds), *La formazione delle parole*. Atti del XXXVII congresso internazionale di studi della Società di Linguistica Italiana. Roma: Bulzoni, 229–247.
- Gaeta, L. (2008). The Copernican revolution of affix ordering. Paper presented at the Workshop on affix ordering in typologically different languages. Vienna, February 3–4.
- Gage, W. W. (1985). Vietnamese in Mon-Khmer perspective. In S. Ratanakul et al. (eds), *Southeast Asian Linguistic Studies Presented to André-G. Haudricourt*. Institute of Language and Culture for Rural Development, Mahidol University, 493–524.
- Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84(3): 474–496.
- Galand, L. (2002). Problématique du nom verbal en berbère. In K. Naït-Zerrad (ed.), *Articles de linguistique berbère. Mémorial Werner Vycichl*. Paris: L'Harmattan, 219–234.
- Galand, L. (2010). *Regards sur le berbère*. Milan: Centro Studi Camito-Semitici.
- Gallego, Á. (2010). Predicados ligeros y valoración de rasgos. *Dicenda* 28: 27–55.
- Gamillscheg, E. (1921). Grundzüge der galloromanischen Wortbildung. In E. Gamillscheg and L. Spitzer (eds), *Beiträge zur romanischen Wortbildungslehre*. Geneva: Olschki, 1–80.
- Ganiev, F. A. (2006). Types of affixes in Turkic Languages. In E. V. Boikova and R. B. Rybakov (eds), *Kinship in the Altaic World: Proceedings of the 48th PIAC*. Weisbaden: Otto Harrassowitz Verlag, 139–44.
- Gardner, A.-C. (2012). Abstract Noun ‘Suffixes’ and Text Type in Old English. In H. Sauer and G. Waxenberger (eds), *English Historical Linguistics 2008. Selected papers from the fifteenth International Conference on English Historical Linguistics (ICEHL 15), Munich, 24–30 August 2008. Volume II: Words, texts and genres*. Amsterdam: Benjamins, 119–132.
- Garrett, A. (2001). Reduplication and infixation in Yurok: morphology, semantics and diachrony. *International Journal of American Linguistics* 67: 264–312.
- Gast, V. and van der Auwera, J. (2006). What is ‘contact-induced grammaticalization’? Evidence from Mayan and Mixe-Zoquean languages. Typescript, University of Antwerp.
- Geeraerts, D. (1993). Vagueness’s puzzles, polysemy’s vagaries. *Cognitive Linguistics* 4: 223–272.

- Geilich, B. (1994). *Nasal-Suffixierung. Eine Studie zur vergleichenden Morphologie der indochinesischen Sprachen*. Münster: Ostasiatisches Seminar.
- Genetti, C., Bartee, E., Coupe, A., Hildebrandt, K., and Lin, Y.-J. (2008). Syntactic aspects of nominalization in five Tibeto-Burman languages of the Himalayan Area. *Linguistics of the Tibeto-Burman Area* 31(2): 97–144.
- Georg, S. (1999/2000). Haupt und Glieder der altaischen Hypothese. Die Körperteilbezeichnungen im Türkischen, Mongolischen und Tungusischen. *Ural-altaische Jahrbücher, neue Folge B* 16: 143–182.
- Georg, S. (2004). Review of *Etymological Dictionary of the Altaic Languages*. *Diachronica* 21(20): 445–450.
- Georg, S. (2005). Reply [to Starostin 2005]. *Diachronica* 22(2): 455–457.
- Georg, S. (2007). *A Descriptive Grammar of Ket. Part I: Introduction, Phonology and Morphology*. Kent, UK: Global Oriental.
- Georg, S., Michalove, P. A., Manaster Ramer, A., and Sidwell, P. J. (1999). Telling general linguists about Altaic. *Journal of Linguistics* 35: 65–98.
- Gerdts, D. B. (2003). The morphosyntax of Halkomelem lexical suffixes. *International Journal of American Linguistics* 69(4): 345–356.
- Gerzenstein, A. (1978). *Lengua Chorote*. Buenos Aires: University of Buenos Aires.
- Gerzenstein, A. (1979). *Lengua Chorote II, Vocabulario*. Buenos Aires: University of Buenos Aires.
- Gerzenstein, A. (1983). *Lengua Chorote. Variedad No. 2*. Buenos Aires: University of Buenos Aires.
- Gerzenstein, A. (1994). *Lengua Maká. Estudio descriptivo*. Buenos Aires: Univsersity of Buenos Aires.
- Gerzenstein, A. (1997a). Chorote word list. In M. R. Key (ed.), *South American Indian Languages*. Computer Database.
- Gerzenstein, A. (1997b). Maka word list. In M. R. Key (ed.), *South American Indian Languages*. Computer Database.
- Gerzenstein, A. (2002). La negación en la lengua maká (mataguayo). Una presentación preliminar. In A. Fernández Garay and L. Golluscio (eds), *Temas de lingüística aborigen II*. Buenos Aires: University of Buenos Aires, 27–52.
- Gerzenstein, A. (2004). Las consonantes laterales y las labializadas en lenguas mataguayas del Chaco argentino paraguayo. In Z. Estrada et al. (eds), *Estudios en lenguas amerindias*. México: University of Sonora, 183–197.
- Ghomeshi, J., Jackendoff, R., Rosen, N., and Russell, K. (2004). Contrastive focus reduplication in English (The Salad-Salad Paper). *Natural Language and Linguistic Theory* 22: 307–357.
- Giakoumaki, E. (2000). *Euphemism. A Linguistic Approach*. Athens: Eleftheria Giakoumak.
- Giegerich, H. (1999). *Lexical Strata in English*. Cambridge: Cambridge University Press.
- Gil, D. (2011). Distributive numerals. In M. S. Dryer and M. Haspelmath (eds), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, ch. 54.
- Gilbers, D. G., Nerbonne, J., and Schaeken, J. (eds) (2000). *Languages in Contact*. (Studies in Slavic and General Linguistics, 28). Amsterdam: Rodopi.
- Gillis, S. (1997). The acquisition of diminutives in Dutch. In W. Dressler (ed.), *Studies in Pre- and Protomorphology*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften, 165–179.
- Gimba, A. M. (2000). Bole verb morphology. PhD dissertation. University of California, Los Angeles.

- Giraud, H. and Grainger, J. (2001). Priming complex words: evidence for supralexical representation of morphology. *Psychonomic Bulletin and Review* 8: 127–131.
- Giraud, H. and Grainger, J. (2003). On the role of derivational affixes in recognizing complex words: Evidence from masked priming. In R. H. Baayen and R. Schreuder (eds), *Morphological Structure in Language Processing*. Berlin: Mouton, 211–234.
- Girault Duvivier, C. (1822), *Grammaire des grammaires, ou analyse raisonnée des meilleurs traités sur la langue française*. 5th edn, Paris: Janet and Cotelte.
- Givón, T. (1993). *English Grammar. A Function-based Introduction. Volume I*. Amsterdam: Benjamins.
- Gleitman, L. R. and Gleitman, H. (1970). *Phrase and Paraphrase: Some Innovative Uses of Language*. New York: W. W. Norton and Co.
- Gleitman, L. R. and Gleitman, H. (1979). Language use and language judgement. In C. J. Fillmore, D. Kempler, and W. S.-Y. Wang (eds), *Individual Differences in Language Ability and Language Behaviour*. New York: Academic Press, 103–126.
- Göksel, A. and Kerslake, C. (2005). *Turkish: A Comprehensive Grammar*. London: Routledge.
- Goldberg, A. E. (1995). *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. (2006). *Constructions at work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Goldenberg, G. (1994). Principles of Semitic word structure. In G. Goldenberg and S. Raz (eds), *Semitic and Cushitic Studies*. Wiesbaden: Harrassowitz, 29–64.
- Goldsmith, J. (1976). *Autosegmental Phonology*. Bloomington, IN: Indiana University Linguistics Club.
- Golla, V. (1970). Hupa grammar. PhD dissertation, Department of Linguistics, University of California Berkeley.
- Golla, V. K. (2007). North America. In Christopher Moseley (ed.), *Encyclopedia of the World's Endangered Languages*. London: Routledge. 1–96.
- Golluscio, L. and Vidal, A. (2009/10). Recorrido sobre las lenguas del Chaco y los aporte a la investigación lingüística, *Amerindia. Les Langues du Chaco* 33/34: 3–40.
- Gomez-Imbert, E. (1996). When animals become ‘rounded’ and ‘feminine’: conceptual categories and linguistic classification in a multilingual setting. In J. Gumperz and S. Levinson (eds), *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press, 438–69.
- Gong Hwang-Cherng 龔煌成 (1977). 古藏文的y及其相關問題 guzangwen de y ji qixiang-guan wenti [Ancient Tibetan y and related questions]. *Bulletin of the Institute of History and Philology, Academia Sinica* 48(2): 205–228. (reprinted in) Gong Hwang-Cherng 龔煌城 (2002). 漢藏語研究論文集 Hanzangyu yanjiulun wenji / Collected Papers on Sino-Tibetan Linguistics. Taipei: 中央研究院語言學研究所籌備處 Zhong yang yan jiu yuan yuyanxue yanjiusuo choubichu, 379–399.
- Gonnerman, L. and Anderson, E. (2001). Graded semantic and phonological similarity effects in morphologically complex words. In S. Bendjaballah, W. Dressler, O. E. Pfeiffer, and M. D. Voikova (eds), *Morphology 2000: Selected Papers from the 9th Morphology Meeting*. Amsterdam: Benjamins, 37–148.
- González, H. A. (2005). A Grammar of Tapiete (Tupi-Guarani). Ph.D. dissertation, University of Pittsburgh.
- Gonzalez-Diaz, V. (2008). *English Adjective Comparison: A Historical Perspective*. Amsterdam: Benjamins.

- Gooch, A. (1970). *Diminutive, Augmentative and Pejorative Suffixes in Modern Spanish*. Oxford: Pergamon Press.
- Good, J. (2003). Strong linearity: three case studies towards a theory of morphosyntactic templatic constructions. Ph.D. dissertation, University of California, Berkeley.
- Good, J. (2006). Rarum begets rarum: a rare clitic and morphosyntactic reduplication in Chechen and Ingush. Rara and Rarissima conference. Max Planck Insititute for Evolutionary Anthropology, Leipzig. Downloadable from <<http://www.acsu.buffalo.edu/~jcgood/jcgood-Rara-Chechen.pdf>>.
- Goudswaard, N. E. (2005). *The Begak (Ida'an) language of Sabah*. Amsterdam: LOT Publications.
- Gràcia, L. and Turon, L. (2000). On appreciative suffixes. *Acta Linguistica Hungarica* 47(1–4): 231–247.
- Gradin, D. (1976). Word affixation in Jeh. *Mon-Khmer Studies* 5: 25–42.
- Gragg, G. B. and Hoberman, R. D. (2012). Semitic. In Z. Frajzyngier and E. Shay (eds), *The Afroasiatic Languages*. Cambridge: Cambridge University Press, 145–235.
- Grainger, J. and Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review* 103: 518–565.
- Grammatika (1980). = D. V. Cygankin (ed.), *Grammatika mordovskih jazykov*. Saransk: Mordovskij gos. universitet.
- Grandi, N. (2005). Sardinian evaluative morphology in typological perspective. In I. Putzu (ed.), *Sardinian in Typological Perspective*. Bochum: Dr. Brockmeyer University Press, 188–209.
- Grandi, N. (2011a). *Evaluative Affixes between Inflection and Derivation: A Typological Survey*. Societas Linguistica Europaea—44th Annual Meeting. Universidad de la Rioja, Logrono. September 10, 2011. Lecture.
- Grandi, N. (2011b). Renewal and innovation in the emergence of Indo-European evaluative morphology. *Lexis* 6: 5–25.
- Grandi, N. and Körtvelyessy, L. (eds). (forthcoming). *Handbook of Evaluative Morphology*. Edinburgh: Edinburgh University Press.
- Graziano-King, J. (1999). Acquisition of comparative forms in English. Ph.D. dissertation, City University of New York.
- Graziano-King, J. and Cairns, H. (2005). The acquisition of English comparatives. *Journal of Child Language* 32(2): 345–373.
- Greenberg, J. (1950). The patterning of root morphemes in Semitic. *Word* 6: 162–181.
- Greenberg, J. H. (1955). *Studies in African Linguistic Classification*. New Haven, CT: Compass Publishing.
- Greenberg, J. H. (1956). The general classification of Central and South American Indian languages. In A. Wallace (ed.), 1960. *Men and Cultures: Selected Papers of the 5th International Congress of Anthropological and Ethnological Sciences*. Philadelphia: University of Pennsylvania, 791–794.
- Greenberg, J. H. (1957). The nature and uses of linguistic typologies. *International Journal of American Linguistics* 23 (2): 68–77.
- Greenberg, J. H. (1963a). *Universals of Language*. Cambridge, MA: MIT Press.
- Greenberg, J. H. (1963b). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (ed.), *Universals of Language*, Cambridge, MA: MIT Press, 73–113.
- Greenberg, J. H. (1963c). *The Languages of Africa*. (International Journal of American Linguistics 29.1, Part II). Bloomington, IN: Indiana University.
- Greenberg, J. H. (1987). *Language in the Americas*. Stanford, CA: Stanford University.

- Greenberg, J. H. (2000–2002). *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*, 2 volumes. Stanford, CA: Stanford University Press.
- Greenhill, S.J., Blust, R., and Gray, R. D. (2008). The Austronesian basic vocabulary database; from bioinformatics to lexomics. *Evolutionary Bioinformatics* 4: 271–283.
- Gregerson, K. (1979). *Predicate and Argument in Rengao Grammar*. Dallas, TX: University of Texas at Arlington, SIL.
- Grégoire, A. (1947). *L'apprentissage du langage. Les deux premières années*. Liège/Paris: Droz.
- Gregová, R. (2009). On phonetic iconicity in evaluative morphology of West Slavonic languages, *Sučasni doslidžennja z inozemnoji filolohiji. Vypusk 7: zbirnyk naukovych prac*. Užhorod: Užhorodskij nacional'nyj universytet, 90–93.
- Gregová, R., Körtvélyessy, L., and Zimmermann, J. (2010). Phonetic iconicity in the evaluative morphology of a sample of Indo-European, Niger-Congo and Austronesian languages. *Word Structure* 3(2): 156–180.
- Grevisse, M. (1980). *Le bon usage. Grammaire française avec des Remarques sur la langue française d'aujourd'hui, onzième édition revue*. Paris/Gembloux: Duculot.
- Grimes, S. (2002). Morphological gemination and root augmentation in three Muskogean languages. MS., Indiana University.
- Grimm, J. (1819–1837). *Deutsche Grammatik*, 4 Tle, Teil 2. 1826. Göttingen: Dieterich'sche Buchhandlung.
- Grimshaw, J. (1990). *Argument Structure*. Cambridge, MA: MIT Press.
- Grimshaw, J. (1993). Semantic structure and semantic content in lexical representation. Unpublished MS. New Brunswick, NJ: Rutgers University.
- Grimshaw, J. (2005). *Words and Structure*. Stanford, CA: CSLI Publications.
- Grönke, U. (1955). *Die Diminutiva des Isländischen. Ein Beitrag zur isländischen Wortkunde*. Göttingen.
- Grossmann, M. and Rainer, F. (2004). *La formazione delle parole in italiano*. Tübingen: Niemeyer.
- Grünthal, R. (2000). Typological characteristics of the Finnic languages: a reappraisal. In J. Laakso (ed.), *Facing Finnic: Some Challenges to Historical and Contact Linguistics*. Castrenianumin toimitteita 59. Helsinki: Helsingin yliopisto, suomalais-ugrilainen laitos, 31–63.
- Grzegorzczak, R., Laskowski, R., and Wróbel, H. (eds) (1999). *Gramatyka współczesnego języka polskiego: Morfologia*. Warszawa: PWN.
- Gualdieri, B. (1998). Mocoví (Guaicuru). Fonologia e morfossintaxe. Ph.D. dissertation, Campinas: University of Campinas.
- Guasch, A. (1983). *El idioma Guaraní. Gramática y antología de prosa y verso*. 6th edn, Asunción: Loyola.
- Guilbert, L. (1975). *La créativité lexicale*. Paris: Librairie Larousse.
- Guilfoyle, E. (2000). Tense and N-Features in Modern Irish. In A. Carnie and E. Guilfoyle (eds), *The Syntax of Verb Initial Languages*. Oxford: Oxford University Press, 61–74.
- Guillaume, A. (2008). *A grammar of Cavineña*. Berlin: Mouton de Gruyter.
- Gurdon, P. R.T. (1914). *The Khasis*. Government of Eastern Bengal and Assam. Available through Project Gutenberg.
- Gurubasave Gowda, K. S. (1975). *Ao Grammar*. Manasagangotri, Mysore: Central Institute of Indian Languages.
- Gustafsson, A. (1979). Treåringars språkliga kreativitet [The linguistic creativity of three-year-olds]. Paper No. 2, Child Language Research Institute, Stockholm University.
- Gvozdev, A. N. (1961). Voprosy izucheniya detskoy rechi [The development of language in children]. Moscow: Akademii Pedagogicheskikh Nauk RSFSR.

- Györke, J. (1934). *Die Wortbildungslehre des Uralischen. (Primäre Wortbildungssuffixe)*. Tartu: Krüger.
- Gyurmé, K. (1992). *Le claire miroir: enseignement de la grammaire tibétaine*. Villeurbanne: Arvillard.
- Haas, M. R. and Swadesh, M. (1932). A journey to the other world. *International Journal of American Linguistics* 7: 195–208.
- Haase, M. (1992). *Sprachkontakt und Sprachwandel im Baskenland: die Einflüsse des Gaskognischen und Französischen auf das Baskische*. Hamburg: Buske.
- Haase, M. (1997). Gascon et basque: bilinguisme et substrat. *Sprachtypologie und Universalienforschung* 50(3): 189–228.
- Haberland, H. (1994). Danish. In E. König and J. van der Auwera (eds), *The Germanic Languages*. London: Routledge, 313–348.
- Hagiwara, H., Sugioka, Y., Ito, T., Kawamura, M., and Shiota, J. (1999). Neurolinguistic evidence for rule-based nominal suffixation. *Language* 75(4): 739–763.
- Hahn, M. (1997). A propos the term gtsug-lag. In H. Krasser, M. T. Much, E. Steinkellner, H. Tauscher (eds), *Proceedings of the 7th Seminar of the International Association for Tibetan Studies*. Wien: Österreichische Akademie der Wissenschaften, 341–348.
- Haig, G. (2001). Linguistic diffusion in present-day East Anatolia: from top to bottom. In A. Y. Aikhenvald and R. M. W. Dixon (eds), *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press, 195–224.
- Haig, G. (2004). Constraints on morpheme repetition in Turkish? In K. Imer and G. Dogan (eds), *Current Research in Turkish Linguistics*. Gazimagusa: Eastern Mediterranean University Press, 3–12.
- Haiman, J. (1974). Concessives, conditionals, and verbs of volition. *Foundations of Language* 11: 341–359.
- Haiman, J. (1997). Repetition and identity. *Lingua* 100: 57–70.
- Hajdú, P. (1987). Die uralischen Sprachen. In P. Hajdú and P. Domokos (eds), *Die uralischen Sprachen und Literaturen*. Budapest: Akadémiai Kiadó, 21–450.
- Hakulinen, A., Vilkkuna, M., Korhonen, R., Koivisto, V., Heinonen, T. R., and Alho, I. (2004). *Iso suomen kielioppi*. [Extensive grammar of Finnish] Helsinki: Suomalaisen Kirjallisuuden Seura. Online version: <<http://scripta.kotus.fi/visk>>.
- Hakulinen, L. (1961). *The structure and development of the Finnish language*. Indiana University Publications, Uralic and Altaic Series, 3. Bloomington, IN: Indiana University. [English translation of an earlier edition of Hakulinen (1979).]
- Hakulinen, L. (1979). *Suomen kielen rakenne ja kehitys*. [The structure and development of the Finnish language]. 4th edn. Helsinki: Otava.
- Hale, K. (1965). Some preliminary observations on Papago morphophonemics. *International Journal of American Linguistics* 31: 295–305.
- Hale, K. L., and Keyser, S. J. (1987). *A View from the Middle* (Lexicon Project Working Papers 10). Cambridge, MA: MIT, Center for Cognitive Science.
- Hale, K. and Keyser, S. J. (eds) (1993). *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press.
- Hale, K. and Keyser, S. J. (1998). The basic elements of argument structure. In *Papers from the UPenn/MIT roundtable on argument structure and aspect*, MITWPL 32: 73–118.
- Hale, K. L., and Keyser, S. J. (2002). *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.
- Hale, K., Laughren, M., and Simpson, J. (1996). Warlpiri Syntax. In J. Jacobs (ed.), *Handbook of Syntax* (Handbücher zur Sprach- und Kommunikationswissenschaft. Berlin: Walter de Gruyter), 1430–1451.

- Hall, C. J. (1992). *Morphology and Mind*. London: Routledge.
- Hall, R. (1956). How we noun-incorporate in English. *American Speech* 31: 83–88.
- Hallap, V. (2000). *Verbaaltuletussufiksids mordva keeltes*. [Verbal derivation suffixes in the Mordvin languages]. Eesti Keele Instituudi Toimetised 5. Tallinn: Eesti Keele Sihtasutus. [Estonian-language version of an unpublished dissertation from 1955.]
- Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry* 4: 3–16.
- Halle, M. and Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (eds), *The View from Building 20*. Cambridge, MA: MIT Press, 111–176.
- Halle, M. and Mohanan, K. P. (1985). Segmental phonology of Modern English. *Linguistic Inquiry* 16: 57–116.
- Haller, F. (2009). Switch-reference in Tibetan. *Linguistics of the Tibeto-Burman Area* 32(2): 45–106.
- Haller, S. and Bartsch, A. J. (2009). Pitfalls in fMRI. *European Radiology* 19(11): 2689–2706.
- Hanique, I. and Ernestus, M. (2012). The role of morphology in acoustic reduction. *Lingue e Linguaggio* 11(2): 147–164.
- Haman, E. (2003). Early productivity in derivation: a case study of diminutives in the acquisition of Polish. *Psychology of Language and Communication* 7: 37–56.
- Haman, E., Zevenbergen, A., Andrus, M., and Chmielewska, M. (2009). Coining compounds and derivations—a crosslinguistic elicitation of word-formation abilities of preschool children and adults in Polish and English. *Polish Psychological Bulletin* 40: 176–192.
- Handel, Z. (2008). What is Sino-Tibetan? Snapshot of a Field and a Language Family in Flux. *Language and Linguistics Compass* 2/3: 422–441.
- Hardy, H. and Montler, T. (1988a). Alabama radical morphology. In W. Shipley (ed.), *In Honor of Mary Haas*. Berlin: Mouton de Gruyter, 377–409.
- Hardy, H. and T. Montler (1988b). Imperfective gemination in Alabama. *International Journal of American Linguistics* 54: 399–415.
- Hargus, S. (2007). *Witsuwit'en Grammar: Phonetics, Phonology and Morphology*. Vancouver: UBC Press.
- Harley, H. (2005). How do verbs get their names? denominal verbs, manner incorporation, and the ontology of verb roots in English. In N. Erteschik-Shir and T. Rapoport (eds), *The Syntax of Aspect*, Oxford: Oxford University Press, 42–64.
- Harley, H. (2009). Compounding in distributed morphology. In R. Lieber and P. Štekauer (eds), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 129–144.
- Harley, H. (2011). On the identity of roots. University of Arizona, ms. <<http://ling.auf.net/lingBuzz/001527>>.
- Harley, H. (2012). Lexical decomposition in modern generative grammar. In W. Hinzen, M. Werning, and E. Mach (eds), *Handbook of Compositionality*. Oxford: Oxford University Press, 328–350.
- Harley, H. and Florez Leyva, M. (2009). Form and meaning in Hiaki (Yaqui) verbal reduplication. *International Journal of American Linguistics* 75: 233–272.
- Harley, H. and Noyer, R. (1999). Distributed morphology. *GLOT International* 4(4): 3–9.
- Harley, H. and Noyer, R. (2000). Licensing in the non-lexicalist lexicon. In B. Peeters (ed.), *The Lexicon/Encyclopedia Interface*. Amsterdam: Elsevier Press, 349–374.
- Harley, H. and Noyer, R. (2003). Distributed morphology. In L. Cheng and R. Sybesma (eds), *The Second Glot International State-of-the-Article Book: The Latest in Linguistics*. Berlin: Mouton de Gruyter, 462–496.

- Harm, M. W. and Seidenberg, M. S. (2004). Computing the meanings of words in reading: co-operative division of labor between visual and phonological processes. *Psychological Review* 111: 662–720.
- Harper, K. (1979). *Suffixes of the Eskimo Dialects of Cumberland Peninsula and North Baffin Island* (Mercury Series No. 54). Ottawa: National Museum of Man.
- Harris, A. (2006). Revisiting anaphoric islands. *Language* 82: 114–130.
- Harris, A. C. and Campbell, L. (1995). *Historical Syntax in Cross-linguistic Perspective*. Cambridge: Cambridge University Press.
- Harris, Z. (1946). From morpheme to utterance. *Language* 22(3): 161–183.
- Harris, Z. (1955). From phoneme to morpheme. *Language* 31(2): 190–222.
- Harrison, S. P. (1973). Reduplication in Micronesian languages. *Oceanic Linguistics* 12: 407–454.
- Harrison, S. P. (1976). *Mokilese Reference Grammar*. Honolulu: The University Press of Hawaii.
- Hartshorne, J. K. and Ullman, M. T. (2006). Why girls say ‘holded’ more than boys. *Developmental Science* 9: 21–32.
- Haspelmath, M. (1990). The grammaticization of passive morphology. *Studies in Language* 14: 25–72.
- Haspelmath, M. (1993). More on the typology of inchoative/causative verb alternations. In B. Comrie and M. Polinsky (eds), *Causatives and Transitivity*. Amsterdam: Benjamins, 87–120.
- Haspelmath, M. (1996). Word-class-changing inflection and morphological theory. In G. Booij and J. van Marle (eds), *Yearbook of Morphology 1995*. Dordrecht: Kluwer, 43–66.
- Haspelmath, M. (1998). How young is Standard Average European? *Language Sciences* 20(3): 271–287.
- Haspelmath, M. (2001). The European linguistic area: Standard Average European. In M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds), *Language Typology and Language Universals: An International Handbook*. Vol. 2. (Handbücher zur Sprach- und Kommunikationswissenschaft, 20.2). New York: Walter de Gruyter, 1492–1510.
- Haspelmath, M. (2002). *Understanding Morphology*. London: Edward Arnold.
- Haspelmath, M. (2003). The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In M. Tomasello (ed.), *The New Psychology of Language*, Vol. 2. Mahwah, NJ: Erlbaum, 211–242.
- Haspelmath, M. (2004). On directionality in language change with particular reference to grammaticalization. In O. Fischer, M. Norde, and H. Perridon (eds), *Up and Down the Cline: The Nature of Grammaticalization* (Typological Studies in Language 59). Amsterdam: Benjamins, 17–44.
- Haspelmath, M. (2009). An empirical test of the agglutination hypothesis. In S. Scalise, E. Magni, and A. Bisetto (eds), *Universals of Language Today* (Studies in Natural Language and Linguistic Theory 76). Dordrecht: Springer, 13–29.
- Haspelmath, M., and Müller-Bardey, T. (2004). Valence change. In C. Lehmann et al. (eds), *Ein internationales Handbuch zur Flexion und Wortbildung* [An International Handbook on Inflection and Word-Formation]. Berlin: de Gruyter Mouton, 1130–1145.
- Haspelmath, M. and Sims, A. (2010). *Understanding Morphology*. 2nd edn. London: Hodder Education.
- Haspelmath, M., König, E., Oesterreicher, W., and Raible, W. (eds) (2001). *Language Typology and Language Universals: An International Handbook*. Vol. 2. (Handbücher zur Sprach- und Kommunikationswissenschaft, 20.2). New York: Walter de Gruyter.

- Haspelmath, M., Dryer, M. S., Gil, D., and Comrie, B. (2005). *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Hathout, N. and Tanguy, L. (2002). Webaffix: finding and validating morphological links on the www. In *Proceedings of the Third International Conference on Language Resources and Evaluation*. Las Palmas de Gran Canaria, Spain, 1799–1804.
- Haugen, E. (1950). The analysis of linguistic borrowing. *Language* 26: 210–231.
- Haugen, J. (2008a). Denominal verbs in Uto-Aztecan. *International Journal of American Linguistics* 74(4): 439–70.
- Haugen, J. (2008b). *Morphology at the Interfaces: Reduplication and Noun Incorporation in Uto-Aztecan*. Amsterdam: Benjamins.
- Haugen, J. (2009). Three challenges for Morphological Doubling Theory. Meertens Instituut, Amsterdam. <<http://www.uni-leipzig.de/~exponet/Slides/Amsterdam/Haugen.pdf>>.
- Hauk, O., Davis, M., Ford, M., Pulvermüller, F., and Marslen-Wilson, W. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage* 30: 1383–1400.
- Haviland, J. (2006). Documenting lexical knowledge. In J. Gippert, N. Himmelmann, and U. Mosel (eds), *Essentials of Language Documentation*. Berlin: Mouton de Gruyter, 129–162.
- Hawkins, J. and Gilligan, G. (1988). Prefixing and suffixing universals in relation to basic word order. *Lingua* 74: 219–259.
- Hay, J. B. (2000). Causes and consequences of word structure. Ph.D. dissertation. Northwestern University.
- Hay, J. B. (2001). Lexical frequency in morphology: is everything relative? *Linguistics* 39(6): 141–170.
- Hay, J. B. (2002). From speech perception to morphology: affix-ordering revisited. *Language* 78: 527–555.
- Hay, J. B. (2003). *Causes and Consequences of Word Structure*. London: Routledge.
- Hay, J. B. and R. H. Baayen. (2002). Parsing and productivity. In G. E. Booij and J. van Marle (eds), *Yearbook of Morphology 2001*. Dordrecht: Kluwer, 203–235.
- Hay, J. B. and Baayen, R. H. (2005). Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences* 9(7): 342–348.
- Hay, J. B. and Plag, I. (2004). What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural Language and Linguistic Theory* 22(3): 565–596.
- Hay, J. B., Kennedy, C. and Levin, B. (1999). Scalar structure underlies telicity in degree achievements. In T. Matthews and D. Strohovitch (eds), *Proceedings of SALT IX*. CLC Publications, 127–144.
- Hayashi, M. (2011). The structure of multiple tenses in Inuktitut. Ph.D. thesis. University of Toronto.
- Hayes, B. and Abad, M. (1989). Reduplication and syllabification in Ilokano. *Lingua* 77: 331–374.
- Hayward, R. J. (2003). Omotic. The ‘empty quarter’ of Afroasiatic linguistics. In J. Lecarme (ed.), *Research in Afroasiatic Grammar II. Selected Papers from the Fifth Conference on Afroasiatic Languages, Paris, 2000*. Amsterdam: Benjamins, 241–262.
- Headley, R. (1977). A Pearic vocabulary. *Mon-Khmer Studies Journal* 6: 69–149.
- Healey, P. M. (1960). *An Agta Grammar*. Manila: Bureau of Printing.
- Heath, J. (1978). Uto-Aztecan _na-class verbs. *International Journal of American Linguistics* 44: 211–22.
- Heath, J. (2005). *A Grammar of Tamashek (Tuareg of Mali)*. Berlin: Mouton de Gruyter.

- Heeschen, V. (1978). *The Mek Languages of Irian Jaya with Special Reference to the Eipo Language*. (grammar_sketch). *Irian* VII: 2.
- Heine, B. (1994). Areal influence on grammaticalization. In Pütz, M. (ed.), *Language Contact and Language Conflict*. Amsterdam: Benjamins, 56–68.
- Heine, B. and Kuteva, T. (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, B. and Kuteva, T. (2003). Contact-induced grammaticalization. *Studies in Language* 27: 529–72.
- Heine, B. and Kuteva, T. (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Heine, B. and Kuteva, T. (2006). *The Changing Languages of Europe*. Oxford: Oxford University Press.
- Heine, B. and Kuteva, T. (2010). Contact and grammaticalization. In R. Hickey (ed.), *The Handbook of Language Contact*. Oxford: Blackwell, 86–105.
- Heine, B. and Miyashita, H. (2008). Accounting for a functional category: German *drohen* ‘to threaten’. *Language Sciences* 30: 53–101.
- Heine, B. and Reh, M. (1984). *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
- Heine, B., Claudi, U., and Hünemeyer, F. (1991). *Grammaticalization. A Conceptual Framework*. Chicago: University of Chicago Press.
- Hellwig, B. (2011). *A Grammar of Goemai*. Berlin: Walter de Gruyter.
- Henderson, J. (2003). The word in Eastern/Central Arrernte. In R. M. W. Dixon and A. Y. Aikhenvald (eds), *Word: A Cross-Linguistic Typology*. Cambridge: Cambridge University Press, 100–124.
- Henderson, J. and Dobson, V. (1994). *Eastern and Central Arrernte to English Dictionary*. Alice Springs: Institute for Aboriginal Development.
- Hengeveld, K. (1992). *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter.
- Henzen, W. ([1947] 1965). *Deutsche Wortbildung*. 3rd edn. Tübingen: Niemeyer.
- Henzen, W. (1947). *Deutsche Wortbildung*. Sammlung kurzer Grammatiken germanischer Dialekte: B, Ergänzungsreihe, 5. Halle: Niemeyer.
- Hercus, L. A. (1982). *The Bāgandji Language*. Canberra: Australian National University.
- Hercus, L. A. (1993). *Paakantyi Dictionary*. Canberra: The Author, with the assistance of the Australian Institute of Aboriginal and Torres Strait Islander Studies.
- Hercus, L. A. (1994). *A Grammar of the Arabana-Wangkangurru Language*, Lake Eyre Basin, South Australia, Series C-128. Canberra: Pacific Linguistics.
- Hetzron, R. (1972). *Ethiopian Semitic: Studies in Classification*. Manchester: Manchester University Press.
- Hetzron, R. (ed.) (1997). *The Semitic Languages*. London: Routledge.
- Hewson, J. (1990). *The Micmac Grammar of Father Pacifique*. Winnipeg: Algonquian and Iroquoian Linguistics.
- Heyvaert, L. (2006). A symbolic approach to deverbal -ee derivation. *Cognitive Linguistics* 17: 337–364.
- Heyvaert, L. (2010). A cognitive-functional perspective on deverbal nominalizations in English. Descriptive findings and theoretical ramifications. In M. Rathert and A. Alexiadou (eds), *The Semantics of Nominalizations across Languages and Frameworks*. Berlin: Mouton de Gruyter, 51–82.
- Hickey, R. (ed.) (2003). *Motives for Language Change*. Cambridge: Cambridge University Press.

- Hilkenbach, L. (2011). Can complexity-based ordering be extended from English to German? Paper presented at the 3rd Vienna Workshop on Affix Order: Advances in Affix Order Research. Vienna, January 15–16.
- Hill, J. (1970). A peeking rule in Cupeño. *Linguistic Inquiry* 1: 534–539.
- Hill, J. H. (2001). Proto-Uto-Aztecan: a community of cultivators in Central Mexico? *American Anthropologist* 103: 913–934.
- Hill, J. (2010). Dating Proto-Uto-Aztecan: New Evidence from Linguistic Paleontology and Lexicostatistics. Paper presented at the annual meeting of the Friends of Uto-Aztecan Conference, Guadalajara, Mexico.
- Hill, J. and Nolasquez, R. (1973). *Mulu'wetum: The First People. Cupeño Oral History and Language*. Banning, CA: Malki Museum Press.
- Hill, J. and Zepeda, O. (1992). Derived words in Tohono O'odham. *International Journal of American Linguistics* 58: 355–404.
- Hill, N. W. (2009). Review of *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. ed. J. A. Matisoff. Berkeley: University of California Press, 2003. [語言暨語言學 *Languages and Linguistics* 10(1): 173–195.]
- Hill, N. W. (2010a). A *Lexicon of Tibetan Verb Stems as Reported by the Grammatical Tradition*. Munich: Bayerische Akademie der Wissenschaften.
- Hill, N. W. (2010b). The converb -las in Old Tibetan. *Bulletin of the School of Oriental and African Studies* 73(2): 245–260.
- Hill, N. W. (2010c). An overview of Old Tibetan synchronic phonology. *Transactions of the Philological Society* 108(2): 110–125.
- Hill, N. W. (2011a). An inventory of Tibetan sound laws. *Journal of the Royal Asiatic Society of Great Britain and Ireland* (3rd Series) 21(4): 441–457.
- Hill, N. W. (2011b). Alternances entrehet b en tibétain ancien et dans les langues tibétaines modernes. *Revue d'Etudes Tibétaines* 20: 115–122.
- Hill, N. W. (2012). Tibetan -las, -nas, and -bas. *Cahiers de Linguistique—Asie Orientale* 41(1): 3–38.
- Hill, N. W. (forthcoming). A note on voicing alternation in the Tibetan verbal system. *Transactions of the Philological Society*.
- Hiltunen, R. (1983). *The Decline of the Prefixes and the Beginnings of the English Phrasal Verb: The Evidence from Some Old and Early Middle English Texts*. Turku: Turun Yliopisto.
- Hoàng, Tuệ, Lý Toàn Thắng, Tạ Văn Thông, et. al. (1986). *Ngữ pháp Tiếng Koho* (A Grammar of Koho). Lam Dong, Vietnam: Sở Văn Hóa và Thông Tin Lâm Đông.
- Hoàng, Văn Hành (1985). *Từ láy trong Tiếng Việt* (Reduplication in Vietnamese). Hanoi: Nhà Xuất Bản Khoa Học Xã Hội.
- Hoàng, Văn Hành (1993). Từ láy trong các ngôn ngữ Katuic ở Việt Nam (Reduplication in the Katuic languages in Vietnam). *Ngôn Ngữ* 4: 8–17.
- Hoàng, Văn Hành (1998). Từ Láy Đôi trong Tiếng Mường (Reduplication in Muong). In *Từ Láy: Những Vấn Đề Còn Để Ngó* (Reduplication: Issues to Consider). Hanoi: Institute of Linguistics, 165–193.
- Hoàng, Văn Ma and Tạ Văn Thông. (1998). *Tiếng Bru-Văn Kiều* (The Bru-Van Kieu language). Hanoi: Nhà Xuất Bản Khoa Học Xã Hội.
- Hockett, C. (1947). Problems of morphemic analysis. *Language* 23: 321–343.
- Hockett, C. (1954). Two models of grammatical description. *Word* 10: 210–231.
- Hockett, C. F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.
- Hockett, C. F. (1994). Conversion: The problem of parts of speech in English. *Etudes Anglaises* 47(2): 171–181.

- Hoffmann, B. (1777). *Lateinische Sprachkunst, zum bessern Gebrauche der Jugend tabellari- sch abgefasset* (6 vols). Bruchsal: Hoffmann (I, VI); Frankenthal: Segel (II); Mannheim: Kurfürstl. Hofdruckerei (III); Karlsruhe: Machlott (IV, V).
- Hoffmann, C. (1963). *A Grammar of the Margi Language*. London: Oxford University Press.
- Hofmann, D. (1961). *Die k-Diminutiva im Nordfriesischen und in verwandten Sprachen*. Köln: Böhlau Verlag.
- Hohenhaus, P. (2004). Identical constituent compounding—a corpus-based study. *Folia Linguistica* 38: 297–331.
- Holden, J. (2012). *Benasní—I Remember. Dene Sułiné Oral Histories with Morphological Analysis*. Leiden: Brill.
- Holes, C. (2004). *Modern Arabic: Structures, Functions, and Varieties*. Rev. edn. Washington, DC: Georgetown University Press.
- Holton, G. (2000). The phonology and morphology of the Tanacross Athabaskan language. Ph.D. dissertation. Department of Linguistics, University of California Santa Barbara.
- Hopper, P. and Traugott, E. C. (2003). *Grammaticalization*. 2nd edn, Cambridge: Cambridge University Press.
- Horecký, J. (1964). *Morfematická štruktúra slovenčiny*. Bratislava: Vydavateľstvo Slovenskej aka- démie vied.
- Horecký, J., Bosák, J., and Buzássyová, K. (1989). *Dynamika slovnej zásoby súčasnej slovenčiny*. Bratislava: Veda.
- Horgan, D. M. (1980). Patterns of variation and interchangeability in some Old English prefixes. *Neuphilologische Mitteilungen*: 127–130.
- Horvath, J., and Siloni, T. (2011). Anticausatives: against reflexivization. *Lingua* 121: 2176–2186.
- Horwood, G. (2001). Anti-faithfulness and subtractive morphology. MS. Rutgers University.
- Howard, P. (1990). *A Dictionary of the Verbs of South Slavey*. Yellowknife, Northwest Territories: Department of Culture and Communications, Government of the Northwest Territories.
- Hudson, J. (1976). Walmatjari. In R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 205–208.
- Huehnergard, J. (2011, 2005). *A Grammar of Akkadian. Harvard Semitic Studies*. Cambridge, MA: Harvard Semitic Museum/Eisenbrauns.
- Huffman, F. E. (1970). *Modern Spoken Cambodian*. Ithaca, NY: Cornell University, Southeast Asia Program.
- Huffman, F. E. (1986). Khmer loanwords in Thai. In R. J. Bickner et al. (eds), *A Conference on Thai Studies In Honor of William J. Gedney*. Center for South and Southeast Asian studies, the University of Michigan, 199–210.
- Hummel, M. (2000). *Adverbale und adverbialisierte Adjektive im Spanischen*. Tübingen: Narr.
- Hupp, J. M., Sloutsky, V. M., and Culicover, P. W. (2009). Evidence for a domain-general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes* 24: 876–909.
- Hurch, B. (1989). Hispanisierung im Baskischen. In N. Boretzky, W. Enninger, and T. Stolz (eds), *Beiträge zum 5. Essener Kolloquium über 'Grammatikalisierung: Natürlichkeit und Systemökonomie'*. Vol. 1 (Essener Beiträge zur Sprachwandelforschung, 7). Bochum: Brockmeyer, 11–35.
- Hurch, B. (ed.) (2005). *Studies on Reduplication*. Berlin: Mouton de Gruyter

- Hurch, B. and Mattes, V. (2005). Über die Entstehung von partieller Reduplikation. In G. Fenk-Oczlon and C. Winkler (eds), *Sprache und Natürlichkeit. Gedenkband für Willi Mayerthaler*. Tübingen: Narr, 137–156.
- Hurford, J. (1990). Nativist and functional explanations in language acquisition. In I. M. Roca (ed.), *Logical Issues in Language Acquisition*. Dordrecht: Foris, 85–136.
- Hutchison, J. (1981). *The Kanuri Language: A Reference Grammar*. Madison, WI: African Studies Program.
- Hyman, L. M. (1979). *Aghem Grammatical Structure: With Special Reference to Noun Classes, Tense-Aspect and Focus Marking*. Los Angeles: Dept. of Linguistics, University of Southern California.
- Hyman, L. M. (2003). Suffix ordering in Bantu: a morphocentric approach. *Yearbook of Morphology* 2002, 245–281.
- Hyman, L. M. (2007). Niger-Congo verb extensions: overview and discussion. In D. L. Payne and J. Peña (eds), *Selected Proceedings of the 37th Annual Conference on African Linguistics*. Somerville, MA: Cascadia Proceedings Project, 149–163.
- Hyman, L. M. (2009). The natural history of verb stem reduplication in Bantu. *Morphology* 19: 177–206.
- Iacobini, C. (2000). Base and direction of derivation. In G. Booij, C. Lehmann, and J. Mugdan, in collaboration with S. Skopeteas (eds), *Morphologie. Morphology. Ein internationales Handbuch zur Flexion und Wortbildung. An international handbook on inflection and word-formation*. 1. Halband. Volume 1. Berlin: Mouton de Gruyter, 865–876.
- Ibanez Moreno, A., and González Torres, E. (2006). Determining the criteria that are useful to identify the derivational and/or compositional capabilities of lexical items in Old English. <http://www.nytud.hu/cescl/proceedings/Moreno-Torres_CESCL.pdf>, accessed February 22, 2013.
- Ikegami, J. (1974). Versuch einer Klassifikation der Tungussischen Sprachen. In G. Hazai and P. Zieme (eds), *Sprache, Geschichte und Kultur der Altaischen Völker*. Berlin: Akademie Verlag, 271–272.
- INDEC (Instituto Nacional De Estadística y Censos). (2004–2005). *Encuesta Complementaria de Pueblos Indígenas*. Publicación online: <http://www.indec.mecon.ar/webcenso/ecpi/index_ecpi.asp>.
- Indo-Aryan Languages (2012). In *Encyclopædia Britannica*. <<http://www.britannica.com/EBchecked/topic/286348/Indo-Aryan-languages>> accessed June 10, 2012.
- Inkelas, S. (1993). Nimboran position class morphology. *Natural Language and Linguistic Theory* 11: 559–624.
- Inkelas, S. (1998). The theoretical status of morphologically conditioned phonology: A case study of dominance effects. *Yearbook of Morphology* 1997: 121–155.
- Inkelas, S. (2011). The interaction between morphology and phonology. In J. Goldsmith, J. Riggle, and A. Yu (eds), *The Handbook of Phonological Theory*. 2nd edn. Malden, MA: Wiley-Blackwell, 68–102.
- Inkelas, S. and Zoll, C. (2005). *Reduplication: Doubling in Morphology*. Cambridge: Cambridge University Press.
- Inkelas, S., Orgun, O. and Zoll, C. (1997). The implications of lexical exceptions for the nature of grammar. In I. Roca (ed.), *Derivations and Constraints in Phonology*. Oxford: Clarendon Press, 393–418.
- Inuit Circumpolar Council. (2010). The Use of the term Inuit in scientific and other circles. 2010. <<http://inuitcircumpolar.com/section.php?ID=25andLang=EnandNav=Section>>.

- Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology* 19: 640–648.
- Ioannidis, J. P. A. and Trikalinos, T. A. (2007). An exploratory test for an excess of significant findings. *Clinical Trials* 4: 245–253.
- Jaberg, K. (1905). Review of Max Roediger. Die Bedeutung des Suffixes -ment. Diss. Berlin 1904, *Archiv für das Studium der neueren Sprachen und Literaturen* 114: 458–462.
- Jackendoff, R. S. (1975). Morphological and Semantic Regularities in the Lexicon. *Language* 51: 639–671.
- Jackendoff, R. (1977). *X-bar Syntax. A Study of Phrase Structure*. Cambridge, MA: MIT Press.
- Jackendoff, R. S. (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.
- Jackendoff, R. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
- Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
- Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jackendoff, R. (2009). Compounding in the parallel architecture and conceptual semantics. In R. Lieber and P. Štekauer (eds), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 105–128.
- Jackendoff, R. (2010). *Meaning and the Lexicon: The Parallel Architecture 1975–2010*. Oxford: Oxford University Press.
- Jackson, H. and Zé Amvela, E. (2007). *Words, Meaning and Vocabulary. An Introduction to Modern English Lexicology*. 2nd edn. London: Continuum.
- Jacob, J. (1968). *Introduction to Cambodian*. London: Oxford University Press.
- Jacobs, R. A. and Rosenbaum, P. S. (eds) (1970). *Readings in English Transformational Grammar*. Waltham, MA: Ginn.
- Jacobson, S. (1984). *Yup'ik Eskimo Dictionary*. Fairbanks: Alaska Native Language Center.
- Jacobson, S. (1995). *A Practical Grammar of the Central Yup'ik Eskimo Language. With Yup'ik Readings written by A. W. Jacobson*. Fairbanks: Alaska Native Language Center.
- Jacobsson, B. (1977). Adverbs, prepositions, and conjunctions in English: A study in gradience. *Studia Linguistica* 31(1): 38–64.
- Jacques, G. (向柏霖) (2003). 嘉絨語、藏語、及上古漢語的-s後綴 Jiarongyu Zangyu ji Shanggu Hanyu de -s houzhui [The -s suffix in Rgyalrong, Tibetan, and Archaic Chinese]. *Minzu Yuwen* 1: 12–15.
- Jacques, G. (2008). 嘉絨語研究 Jiarongyu Yanjiu [A Study of Rgyalrong]. Beijing: Nationalities Press.
- Jacques, G. (2010a). The origin of the reflexive prefix in Rgyalrong languages. *Bulletin of the School of Oriental and African Studies* 73(2): 261–268.
- Jacques, G. (2010b). A possible trace of verb agreement in Tibetan. *Himalayan Linguistics Journal* 9(1): 41–49.
- Jacques, G. (2012a). Argument demotion in Japhug Rgyalrong. In G. Authier and K. Haude (eds), *Ergativity, Valency and Voice*. Berlin: De Gruyter Mouton, 199–225.
- Jacques, G. (2012b). Agreement Morphology: The Case of Rgyalrongic and Kiranti. *Language and Linguistics* 13(1): 83–116.
- Jacques, G. (2012c). From denominal derivation to incorporation. *Lingua* 122: 1207–1231.
- Jacques, G. (2012d). An internal reconstruction of Tibetan stem alternations. *Transactions of the Philological Society* 110(2): 212–224.
- Jacques, G. (forthcoming). The genetic position of Chinese. CRNS/CRLAO. In R. Sybesma, J. Huang, W. Behr, and Z. Handel, (eds), *Encyclopedia of Chinese Languages and Linguistics*.

- <http://cnrs.academia.edu/GuillaumeJacques/Papers/1009403/The_genetic_position_of_Chinese> accessed August 15, 2012.
- Jacques, G. (向柏霖) and Chen Zhen (2004). 茶堡話的重疊形式 Chabaohua de chongdie xingshi [Reduplicated forms in Japhug]. *Minzu Yuwen* 4: 7–11.
- Jaggar, P. (2001). *Hausa*. Amsterdam: Benjamins.
- Jakobi, A., and Crass, J. (2004). *Grammaire du Beria: langue saharienne*. Cologne: Rüdiger Köppe.
- Jakubowski, P. (2006). On derivational suffixes in three late Middle English romances: Guy of Warwick, Bevis of Hampton, and Sultan of Babylon. *Studia Anglica Posnaniensia: An International Review of English Studies* 42: 189–201.
- Janda, L. A. (2011). Metonymy in word-formation. *Cognitive Linguistics* 22(2), 359–392.
- Janda, L. A. and Townsend, C. E. (2000). *Czech*. Munich: Lincom Europa.
- Janhunen, J. (ed.) (2003). *The Mongolic Languages*. London: Routledge.
- Janhunen, J. (2006). Mongolic languages. In K. Brown (ed.), *The Encyclopedia of Language and Linguistics*. Amsterdam: Elsevier, 231–234.
- Janssen, N., Bi, Y., and Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access for Mandarin Chinese and English compounds. *Language and Cognitive Processes* 23(7–8): 1191–1223.
- Jarceva, V. (ed.) (1997). *Jazyki mira. Mongol'skie, tunguso-manchzhurskie jazyki, japonskij jazyk, korejskij jazyk* [Languages of the world: Mongolic languages, Tungusic languages, Japanese, Korean]. Moscow: Indrik.
- Jarmulowicz, L., Taran, V. L., and Hay, S. E. (2008). Lexical frequency and third-graders' stress accuracy in derived English word production. *Applied Psycholinguistics* 29: 213–235.
- Järvikivi, J. and Niemi, J. (2002). Allomorphs as paradigm indices: on-line experiments with Finnish free and bound stems. *SKY Journal of Linguistics* 15: 119–143.
- Järvikivi, J. and Pyykkönen, P. (2011). Sub- and supralexic information in early phases of lexical access. *Frontiers in psychology*, 2. doi: 10.3389/fpsyg.2011.00282.
- Järvikivi, J., Bertram, R., and Niemi, J. (2006). Affixal salience and the processing of derivational morphology: the role of suffix allomorphy. *Language and Cognitive Processes* 21(4): 394–431.
- Jastrow, O. (2005). Uzbekistan Arabic: A language created by Semitic-Iranian-Turkic linguistic convergence. In E. A. Csató, B. Isaksson, and C. Jahani (eds), *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*. London: Routledge Curzon, 133–139.
- Jeddore, R. (1976). *Labrador Inuit Uqausingit*. Nain: the Labrador Inuit Committee on Literacy.
- Jelinek, E. and Demers, R. (1994). Predicates and pronominal arguments in Straits Salish. *Language* 70: 697–736.
- Jelinek, E. and Escalante, F. (1988). 'Verbless' possessive sentences in Yaqui. In W. Shipley (ed.), *In Honor of Mary Haas: Papers from the Haas Festival Conference on Native American Linguistics*, Berlin: Mouton de Gruyter, 411–29.
- Jendraschek, G. (2006). Basque in contact with Romance languages. In A. Y. Aikhenvald and R. M. W. Dixon (eds), *Grammars in Contact: A Cross-Linguistic Typology*. Oxford: Oxford University Press, 143–162.
- Jenkins, R. S. (2002). *Language Contact and Composite Structures in New Ireland, Papua New Guinea*. Ann Arbor: UMI Dissertation Services.
- Jenner, P. and Saverous Pou (1982). A lexicon of Khmer morphology. *Mon-Khmer Studies* 9–10: 1–517.
- Jenny, M. (2003). New infixes in Spoken Mon. *Mon-Khmer Studies* 33: 183–194.

- Jenny, M. (2005). *The verb system of Mon.* Arbeiten des Seminars für Allgemeine Sprachwissenschaft 19. University of Zurich.
- Jensen, J. T. (1990). *Morphology. Word Structure in Generative Grammar*. Amsterdam: Benjamins.
- Jescheniak, J. D. and Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition* 20(4): 824–843.
- Jespersen, O. (1942). *A Modern English Grammar on Historical Principles*. Part VI: *Morphology*. Copenhagen: Einar Munksgaard.
- Jetté, J. and Jones, E. (2000). *Koyukon Dictionary*. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks.
- Ježek, E. and Melloni, C. (2011). Nominals, polysemy and co-predication. MS. University of Pavia and University of Verona.
- Ji, H., Gagné, C. and Spalding, T. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language* 65: 406–430.
- Jin, P., Tan K., Qu A., Lin X. (1958). 嘉戎語梭磨話的語音及形態（續） Jiarongyu Suomohua de yuyin ji xingtai (xu) [Phonology and morphology in Suomo Rgyalrong, Part II], *Yuyan Yanjiu* 3: 71–108.
- Joanisse, M. F. and Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences* 96: 7592–7597.
- Johanson, L. (1992). Strukturelle Faktoren in türkischen Sprachkontakten. In Sitzungsberichte der Wissenschaftlichen Gesellschaft an der Johann Wolfgang Goethe-Universität Frankfurt am Main, vol. 29(5). Stuttgart: Franz Steiner, 169–299.
- Johanson, L. (2002a). *Structural Factors in Turkic Language Contacts*. London: Curzon.
- Johanson, L. (2002b). Contact-induced linguistic change in a code-copying framework. In Jones, M. C. and E. Esch (eds), *Language Change: The Interplay of Internal, External and Extra-Linguistic Factors*. (Contributions to the Sociology of Language, 86). Berlin: Mouton: de Gruyter, 285–313.
- Johanson, L. and Csató, Á. (1998). *The Turkic Languages*. London: Routledge.
- Johns, A. (1992). Deriving ergativity. *Linguistic Inquiry* 23: 57–88.
- Johns, A. (1999). On the lexical semantics of affixal meaning ‘want’ in Inuktitut. *International Journal of American Linguistics* 65(2): 176–200.
- Johns, A. (2007). Restricting noun incorporation: root movement. *Natural Language and Linguistic Theory* 25: 535–576.
- Johns, A. (2009). Additional facts about noun incorporation (in Inuktitut). *Lingua* 119(2): 185–198.
- Johns, A. (2010). Word internal nominalization in Inuktitut. Paper delivered at the LSA in Baltimore.
- Jones, B. (2011). *A Grammar of Wangkajunga: A Language of the Great Sandy Desert of North Western Australia*. Canberra: Pacific Linguistics.
- Jong, F. de. (2007). *A Grammar of Modern Uyghur*. Utrecht: Houtsma.
- Jungraithmayr, H. (1963). Die Sprache der Sura (Mwaghavul) in Nordnigerien. *Afrika und Übersee* 47: 8–89.
- Jungraithmayr, H. (1990). *Lexique Mokilko*. Berlin: Reimer.
- Jungraithmayr, H. and Adams, A. (1992). *Lexique Migama*. Berlin: Reimer.

- Jurafsky, D. (1993). Universals in the semantics of the diminutive, Parasession on semantic typology and semantic universals. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 19(1): 423–436.
- Jurafsky, D. (1996). Universal tendencies in the semantics of the diminutive. *Language* 72(3): 533–578.
- Kachru, Y. (2006). *Hindi*. Amsterdam: Benjamins.
- Kačmárová, A. (2010). A Short Survey of Diminutives in Slovak and English, *English Matters*, Prešov: PU. 17–21. Retrieved from <<http://www.pulib.sk/elpub2/FHPV/Kacmarova1/index.html>>.
- Kadazan Dusun Cultural Association (1995). *Kadazan Dusun-Malay-English Dictionary*. Kota Kinabalu: Sabah.
- Kamp, H. and Rossdeutsch, A. (1994). Remarks on lexical structure and DRS construction. *Theoretical Linguistics* 20: 97–164.
- Kandel, S., Spinelli, E., Tremblay, A., Guerassimovitch, H., and Álvarez, C. J. (2012). Processing prefixes and suffixes in handwriting production. *Acta psychologica* 140(3): 187–195.
- Kangasmaa-Minn, E. (1987). Functional derivation in Finno-Ugric. In *Papers on Derivation in Uralic. Szegeder und Turkuur Beiträge zur uralischen Derivation*. Szeged: Univ. Szeged de Attila József Nominata, 5–27.
- Kangasmaa-Minn, E. (1992). Pienimuotoinen kieliperimämme [English summary: Our unobtrusive linguistic inheritance]. *Sananjalka* 34: 21–31.
- Kaper, W. (1959). *Kindersprachforschung mit Hilfe des Kindes*. Groningen: Wolters.
- Kaplan, L. (1981). *Phonological Issues in North Alaskan Inupiaq*. Alaska Native Language Center research papers, no. 6. Fairbanks, Alaska (Alaska Native Language Center, University of Alaska, Fairbanks 99701): Alaska Native Language Center.
- Kari, J. (1979). *Athabaskan Verb Theme Categories: Ahtna*. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks.
- Kari, J. (1989). Affix positions and zones in the Athapaskan verb complex: Ahtna and Navajo. *International Journal of American Linguistics* 55: 424–454.
- Kari, J. (1990). *Ahtna Athabaskan Dictionary*. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks.
- Kari, J. (1992). Some concepts in Ahtna Athabaskan word formation. In M. Aronoff (ed.), *Morphology Now*, Albany, NY: State University of New York Press, 107–131.
- Karmanova, A. N. (2002). *Torjakyvjas tečasnog da artmannog*. [Structure and formation of words]. Syktyvkar: Syktyvkarskij gosudarstvennyj universitet.
- Karnchana, N. (1978). The syllabic and morphological structure of Cambodian words. *Mon-Khmer Studies* 7: 183–200.
- Kasik, R. (1997). Typology of Estonian and Finnish word-formation: the verb. In M. Erelt (ed.), *Estonian. Typological Studies II*. Tartu: Tartu University Press, 43–72.
- Kastovsky, D. (1969). Wortbildung und Nullmorphem. *Linguistische Berichte* 2: 1–13.
- Kastovsky, D. (1980). Zero in morphology: a means of making up for phonological losses? In J. Fisiak (eds), *Historical Morphology*. The Hague: Mouton, 213–250.
- Kastovsky, D. (1982). *Wortbildung und Semantik*. Tübingen/Düsseldorf: Francke/Bagel.
- Kastovsky, D. (1986). Productivity in word formation. *Linguistics* 24: 585–600.
- Kastovsky, D. (1992a). The formats change—the problems remain: word-formation theory between 1960 and 1990. In M. Pütz (ed.), *Thirty Years of Linguistic Evolution. Studies in Honour of René Dirven on the Occasion of His Sixtieth Birthday*. Amsterdam: Benjamins, 28–310.

- Kastovsky, D. (1992b). Semantics and Vocabulary. In R. M. Hogg (ed.), *The Cambridge History of the English Language*. Cambridge: Cambridge University Press, 290–408.
- Kastovsky, D. (1994). Verbal derivation in English: A historical survey or Much Ado about Nothing. In D. Britton (ed.), *English Historical Linguistics 1994. Papers from the 8th International Conference on English Historical Linguistics*. Amsterdam: Benjamins, 93–117.
- Kastovsky, D. (1996). The place of word-formation in grammar. In K. R. Jankowsky (ed.), *Multiple Perspectives on the Historical Dimensions of Language*. Münster: Nodus, 227–243.
- Kastovsky, D. (1997). Sign-oriented vs. form-oriented linguistics and word-formation. *Studia Anglica Posnaniensia* XXXI: 79–90.
- Kastovsky, D. (2000). Words and word-formation: morphology in OED. In L. Muggleston (ed.), *Lexicography and the OED. Pioneers in the Untrodden Forest*. Oxford: Oxford University Press, 110–125.
- Kastovsky, D. (2002). The derivation of ornative, locative, ablative, privative and reversative verbs in English: a historical sketch. In T. Fanego, M. J. López-Couso, and J. Pérez-Guerra (eds), *English Historical Syntax and Morphology* (Amsterdam Studies in the Theory and History of Linguistic Science IV: Current Issues in Linguistic Theory: 223). Amsterdam: Benjamins, 99–109.
- Kastovsky, D. (2005). Conversion and/or zero: word-formation theory, historical linguistics and typology. In L. Bauer and S. Valera (eds), *Approaches to Conversion/Zero-derivation*. Münster: Waxmann, 31–49.
- Kastovsky, D. (2006a). Historical morphology from a typological point of view: examples from English. In T. Nevalainen, J. Klemola, and M. Laitinen (eds), *Types of Variation: Diachronic, Dialectal and Typological Interfaces*. Amsterdam: Benjamins, 53–80.
- Kastovsky, D. (2006b). Morphology as word-formation in the 20th-century linguistics: a survey. In S. Aourou, H.-J. Koerner, and K. V. Niederehe (eds), *History of the Language Sciences. An International Handbook on the evolution of the Study of Language from the Beginnings to the Present*. vol. 3. Berlin: Walter de Gruyter.
- Kastovsky, D. (2006c). Typological changes in derivational morphology. In A. van Kemenade and B. Los (eds), *The Handbook of the History of English*. Malden, MA: Blackwell, 151–176.
- Kastovsky, D. (2010). Translation techniques in the terminology of Ælfric's Grammar: semantic loans, loan translations and word-formation. In M. Kytö et al. (eds), *Language Change and Variation from Old English to Late Modern English* (Linguistic Insights: Studies in Language and Communication: 114). Bern: Peter Lang, 163–174.
- Kastovsky, D. (2012). Morphology. In A. Bergs and L. Brinton. *English Historical Linguistics, Vol. I, Part II*, 129–148. [Handbücher zur Sprach- und Kommunikationswissenschaft 34.1].
- Kastovsky, D. (2013). *A Historical Morphology of English*. Edinburgh textbooks on the English language—advanced. Edinburgh: Edinburgh University Press.
- Kaufman, D. (2003). Paradigm effects and the affix-shape/position generalization. In G. Garding and M. Tsujimura, (eds), *WCCFL 22 Proceedings*. Somerville, MA: Cascadia Press, 273–286.
- Kaufman, D. (2009). Austronesian nominalism and its consequences: a Tagalog case study. *Theoretical Linguistics* 35: 1–49.
- Kaufman, T. (1974a). Meso-American Indian languages. *Encyclopaedia Britannica*, 15th edn, ed. by P.W. Goetz, 11. Chicago, 956–63.
- Kaufman, T. (1974b). *Idiomas de Mesoamerica* (Seminario de Integración Social Guatemalteca, publicación no. 33). Guatemala: J. de Pineda Ibarra.

- Kaufman, T. (1990). Language history in South America: What we know and how to know more. In D. Payne (ed.), *Amazonian Linguistics. Studies in Lowland South American Languages*. Austin: University of Texas, 13–73.
- Kaunisto, M. (2001). Nobility in the history of adjectives ending in -ic and -ical. *Linguistic Association of Canada and the United States Forum* 27: 35–45.
- Kaunisto, M. (2004). Adjective Pairs ending in -ic and -ical: A study of variation and change across centuries. *Dissertation Abstracts International, Section C: Worldwide* 65(3): 585.
- Kay, P. and Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: the what's X doing Y? Construction. *Language* 1–33.
- Kaye, A. S. (ed.). (2007). *Morphologies of Asia and Africa*, vol. 1. Winona Lake, Indiana: Eisenbrauns.
- Kazanina, N. (2011). Decomposition of prefixed words in Russian. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(6): 1371.
- Kečkéšová, V. (2012). Sound symbolism and second language acquisition. MA thesis. P. J. Šafárik University, Košice.
- Keegan, J. M. (1997). *A Reference Grammar of Mbay*. Munich: LINCOM Europa.
- Kemmer, S. (1992). Grammatical prototypes and competing motivations in a theory of linguistic change. In G. W. Davis and G. K. Iverson (eds), *Explanation in Historical Linguistics*. Amsterdam: Benjamins, 145–166.
- Kempe, V. and Brooks, P. J. (2001). The role of diminutives in the acquisition of Russian gender: Can elements of child-directed speech aid in learning morphology? *Language Learning* 51: 221–256.
- Kempe, V., Brooks, P. J., Mironova, N., and Fedorova, O. (2003). Diminutivization supports gender acquisition in Russian children. *Journal of Child Language* 30: 471–485.
- Kemps, R., Ernestus, M., Schreuder, R., and Baayen, R. H. (2004). Processing reduced word forms: The suffix restoration effect. *Brain and Language* 19: 117–127.
- Kemps, R., Ernestus, M., Schreuder, R., and Baayen, R. H. (2005a). Prosodic cues for morphological complexity: the case of Dutch noun plurals. *Memory and Cognition* 33: 430–446.
- Kemps, R., Wurm, L. H., Ernestus, M., Schreuder, R., and Baayen, R. H. (2005b). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes* 20: 43–73.
- Kennedy, C. and McNally, L. (2005). Scale structure and the semantic typology of gradable predicates. *Language* 81: 345–381.
- Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.
- Kerleroux, F. (1999). Identification d'un procédé morphologique. La conversion. *Faits de Langues: Revue de Linguistique* 14: 89–100.
- Kernan, K. T. and Blount, B. G. (1966). The acquisition of Spanish grammar by Mexican children. *Anthropological Linguistics* 8: 1–14.
- Kertz, L. (2006). Evaluative adjectives: an adjunct control analysis. In D. Baumer, D. Montero, and M. Scanlon (eds), *WCCFL 28 Proceedings*. Somerville, Cascadia Press, 229–235.
- Keszler, B. (2006). *Magyar Grammatika*. Budapest: Nemzeti Tankönyvkiadó.
- Keuleers, E. and Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* 42(3): 627–633.
- Keuleers, E., Diependaele, K., and Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: a lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology*, 1. doi: 10.3389/fpsyg.2010.00174.

- Keuleers, E., Lacey, P., Rastle, K., and Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods* 44(1): 287–304.
- Keune, K., Ernestus, M., Van Hout, R., and Baayen, R. H. (2005). Social, geographical, and register variation in Dutch: From written ‘mogelijk’ to spoken ‘mok’. *Corpus Linguistics and Linguistic Theory* 1: 183–223.
- Keune, K., van Hout, R., and Baayen, R. H. (2006). Socio-geographic variation in morphological productivity in spoken Dutch: a comparison of statistical techniques. In J.-M. Viprey (ed.), *Actes des 8es journées internationales d’analyse statistique des données textuelles*, vol. 2, Presses Universitaires de Franche-Comté, 571–580.
- Keune, K., van Hout, R. and Baayen, R. H. Derivational and lexical productivity across written and spoken Dutch. Manuscript.
- Key, H. (1965). Some semantic functions of reduplication in various languages. *Anthropological Linguistics* 7: 88–101.
- Keyser, S. J. and Roeper, T. (1984). On the middle and ergative constructions in English. *Linguistic Inquiry* 15: 381–416.
- Kibrik, A. (1995/2005). Inflection vs. derivation and the template for Athabaskan verb morphology. Manuscript.
- Kibrik, A. (2012). Toward a typology of verbal lexical systems: A case study in Northern Athabaskan. *Linguistics*, 50(3): 495–532.
- Kiefer, F. (2003). Alaktan [Morphology]. In K. É. Kiss, F. Kiefer and P. Siptár (eds), *Új magyar nyelvtan* [New Hungarian Grammar]. Budapest: Osiris, 189–284.
- Kiefer, F. (2010). Areal-typological aspects of word-formation: the case of aktionsart-formation in German, Hungarian, Slavic, Baltic, Romani and Yiddish. In F. Rainer, W. U. Dressler, D. Kastovsky, and H. C. Luschützky (eds), *Variation and Change in Morphology. Selected Papers from the International Morphology Meeting, Vienna, February 2008*. Amsterdam: Benjamins, 129–147.
- Kiefer, F. (forthcoming). Hungarian. In P. O. Müller, I. Ohnheiser, S. Olsen, and F. Rainer (eds), *Word-Formation. An International Handbook of Languages in Europe*. Berlin: De Gruyter Mouton.
- Kilham, C., Pamulkan, M., Pootchemunka, J., and Wolmby, T. (1986). *Dictionary and Source Book of the Wik-Mungkan language*. Darwin: SIL-AAB.
- Kimball, G. D. (1991). *Koasati Grammar*. Lincoln, NE: University of Nebraska Press.
- Kimenyi, A. (2002). *A Tonal Grammar of Kinyarwanda: An Autosegmental and Metrical Analysis*. Lewiston, NY: E. Mellen Press.
- Kimura, D. (2000). *Sex and Cognition*. Cambridge, MA: The MIT Press.
- King, K. and Melzi, G. (2004). Intimacy, imitation and language learning: Spanish diminutives in mother-child conversation. *First Language* 24: 241–261.
- Kiparsky, P. (1982a). Word-formation and the lexicon. In F. Ingemann (ed.), *1982 Mid-America Linguistics Conference Papers*. Lawrence, Kansas: Department of Linguistics, University of Kansas, 1–29.
- Kiparsky, P. (1982b). Lexical morphology and phonology. In I. S. Yang (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing, 1–91.
- Kiparsky, P. (1982c). From cyclic phonology to lexical phonology. In H. van der Hulst and N. Smith (eds), *The Structure of Phonological Representation. Volume 1*. Dordrecht: Foris, 131–176.
- Kiparsky, P. (1986). The phonology of reduplication. Unpublished MS., Stanford University.

- Kiparsky, P. (1996). Allomorphy or morphophonology? In R. Singh (ed.), *Trubetzkoy's Orphan: Proceedings of the Montreal Roundtable 'Morphophonology: Contemporary Responses'*. Amsterdam: Benjamins, 13–31.
- Kiparsky, P. (1997). Remarks on denominal verbs. In A. Alsina, J. Bresnan, and P. Sells (eds), *Argument Structure*. Stanford, CA: CSLI Publications, 473–499.
- Kiparsky, P. (2001). Structural case in Finnish. *Lingua* 111: 315–376.
- Kiyomi, S. (1995). A new approach to reduplication: a semantic study of noun and verb reduplication in the Malayo-Polynesian languages. *Linguistics* 33: 1145–1167.
- Klamer, M. (1998). *A grammar of Kambara*. Berlin: Mouton de Gruyter.
- Kleiber, G. (2010). Petit essai pour montrer que la polysémie n'est pas un ... sens interdit, *Cahiers de lexicologie* 96: 129–149.
- Klein, D. E. and Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language* 45: 259–282.
- Kleinschmidt, S. (1851). *Grammatik der Grönländischen Sprache*. Reprinted 1968. Berlin: George Olms.
- Klembárová, E. (2012). *Contrastive Analysis of Word-Formation in Children of Different Age*. MA thesis. P. J. Šafárik University, Košice.
- Kliegl, R., Dambacher, M., Dimigen, O., Jacobs, A. M., and Sommer, W. (2012). Eye movements and brain electric potentials during reading. *Psychological Research* 76(2): 145–158.
- Klingvall, E. (2008). (De)composing the middle: a minimalist approach to middles in English and Swedish. Ph.D. dissertation. Lund University.
- Kneisl, M. (1978). *Die Verbalbildung im Syrjänischen*. München: Traut.
- Knuth, D. E. (1973). *The Art of Computer Programming*. Vol. 3: *Sorting and Searching*. Reading, MA: Addison-Wesley.
- Koch, H. (1984). The category of 'associated motion' in Kaytej. *Language in Central Australia* 1: 23–34.
- Koch, H. (2003). Morphological reconstruction as an etymological method. In B. J. Blake, K. Burridge, and J. Taylor (eds), *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics*, Melbourne, August 13–17, 2001. Amsterdam: Benjamins, 271–291.
- Koch, H. (2007). An overview of Australian traditional languages. In G. Leitner and I. G. Malcolm (eds), *The Habitat of Australia's Aboriginal Languages: Past, Present and Future* (Trends in Linguistics, Studies and Monographs 179). Berlin: Mouton de Gruyter, 23–56.
- Koch, H. and Hercus, L. (2013). Obscure vs transparent cognates in linguistic reconstruction. In R. Mailhammer (ed.), *Lexical and Structural Etymology: Beyond Word Histories*. Berlin: Walter De Gruyter, 33–52.
- Koch, K., and Matthewson, L. (2009). The lexical category debate in salish and its relevance for Tagalog. *Theoretical Linguistics* 35: 125–137.
- Koch, P. (2001). Bedeutungswandel und Bezeichnungswandel. Von der kognitiven Semasiologie zur kognitiven Onomasiologie. *Zeitschrift für Literaturwissenschaft und Linguistik* 121: 7–36.
- Koester, D. and Schiller, N. O. (2011). The functional neuroanatomy of morphology in language production. *NeuroImage* 55(2): 732–741.
- Koontz-Garboden, A. (2005). On the typology of state/change of state alternations. *Yearbook of Morphology* 2005: 83–117.
- Koontz-Garboden, A. (2006). The states in changes of state. In press. In Proceedings of the 32nd annual meeting of the Berkeley Linguistics Society.

- Koontz-Garboden, A. (2007). Aspectual coercion and the typology of change of state predicates. *Journal of Linguistics* 43: 115–152.
- Koontz-Garboden, A. (2009a). Anticausativization. *Natural Language and Linguistic Theory* 27: 77–138.
- Koontz-Garboden, A. (2009b). Ulwa verb class morphology. *International Journal of American Linguistics* 75: 453–512.
- Koontz-Garboden, A. (2010). The lexical semantics of derived statives. *Linguistics and Philosophy* 33: 285–324.
- Koontz-Garboden, A. (2011). The roots of change of state verbs. Paper presented at Approaches to the Lexicon (Roots III), Jerusalem, June 14, 2011.
- Koontz-Garboden, A. (2012a). The Monotonicity Hypothesis. In L. McNally and V. Demonte (eds), *Telicity, Change, and State: A Cross-categorical View of Event Structure*. Oxford: Oxford University Press, 139–161.
- Koontz-Garboden, A. (2012b). The universality of lexical categories: comments on Chung. *Theoretical Linguistics* 38: 103–117.
- Koontz-Garboden, A. (forthcoming). Thoughts on diagnosing morphomicity: a case study from Ulwa. In A. Luís and R. Bermúdez-Otero (eds), *The Morphome Debate: Diagnosing and Analyzing Morphomic Patterns*. Oxford: Oxford University Press.
- Kornexl, L. (2008). Women and other ‘small things’: -Ette as a feminine marker. In R. Dury, M. Gotti, and M. Dossena (eds), *English Historical Linguistics 2006, Volume II: Lexical and Semantic Change* (Amsterdam Studies in the Theory and History of Linguistic Science IV: Current Issues in Linguistic Theory: 296). Amsterdam: Benjamins, 241–257.
- Kornfeld, L. M. (2009). IE, Romance: Spanish. In R. Lieber and P. Štekauer (eds), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 436–452.
- Kornfilt, J. (1997). *Turkish*. London: Routledge.
- Körtvélyessy, L. (2009). Productivity and Creativity in Word-Formation. *Onomasiology Online* 10(1): 1–22.
- Körtvélyessy, L. (2010). *Vplyv sociolingvistických faktorov na produktivitu v slovotvorbe*. Slovacontact: Prešov.
- Körtvélyessy, L. (2012). *Evaluative Morphology from Cross-linguistic Perspective*. Budapest: Habilitationsschrift. ELTE.
- Kossmann, M. (2007). Berber. In A. S. Kaye (ed.), *Morphologies of Asia and Africa*, vol. 1. Winona Lake, IN: Eisenbrauns, 429–446.
- Kossmann, M. (2012). Berber. In Z. Frajzyngier, and E. Shay (eds), *The Afroasiatic Languages*. Cambridge: Cambridge University Press, 18–101.
- Kostov, K. (1973). Zur Bedeutung des Zigeunerischen für die Erforschung grammatischer Interferenzerscheinungen. *Linguistique Balkanique* 16(2): 99–113.
- Koul, O. N. (2008). *Modern Hindi Grammar*. Springfield VA: Dunwoody Press.
- Koziol, H. (1972). *Handbuch der englischen Wortbildungslehre*. Heidelberg: Winter.
- Kratschmer, A. (1996). Homonymie. In G. Ueding (ed.), *Historisches Wörterbuch der Rhetorik*, vol. 3. Tübingen: Niemeyer, 1535–1542.
- Kratzer, A. (1995). Stage-Level and Individual-Level Predicates. In G. N. Carlson and F. J. Pelletier (eds), *The Generic Book*. Chicago: University of Chicago Press, 125–175.
- Kratzer, A. (1996). Severing the external argument from its verb. In J. Rooryck and L. Zaring (eds), *Phrase Structure and the Lexicon*. Dordrecht, Kluwer, 109–137.
- Kratzer, A. (2000). Building statives. *Berkeley Linguistic Society* 26: 385–399.

- Krause, S. (1980). Topics in Chukchee phonology and morphology. Ph.D. thesis. University of Illinois Urbana-Champaign.
- Krauss, M. E. (2000). Koyukon kinship. In J. Jetté and E. Jones (J. Kari, editor-in-chief), *Koyukon Athabaskan dictionary*. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks, 815–822.
- Krejnovich, E. A. (1968). *Glagol ketskogo jazyka [The Ket verb]*. Leningrad: Nauka.
- Kroeger, P. (1993). *Phrase Structure and Grammatical Relations in Tagalog*. Stanford, CA: CSLI.
- Krott, A. (2001). Analogy in morphology: the selection of linking elements in Dutch compounds. Ph.D. dissertation. University of Nijmegen.
- Krott, A., Schreuder, R., and Baayen, R. H. (1999). Complex words in complex words. *Linguistics* 37(5): 905–926.
- Kruspe, N. (2004). *A Grammar of Semelai*. Cambridge: Cambridge University Press.
- Kryuchkova, T., Tucker, B. V., Wurm, L., and Baayen, R. H. (2012). Danger and usefulness in auditory lexical processing: evidence from electroencephalography. *Brain and Language* 122: 81–91.
- Kuczaj, S. A., II. (1979). Evidence for a language learning strategy: on the relative ease of acquisition of prefixes and suffixes. *Child Development* 50: 1–13.
- Kuipers, A. H. (1974). *The Shuswap Language: Grammar, Texts, Dictionary*. The Hague: Mouton.
- Kuperman, V. and Van Dyke, J. (2011). Individual differences in visual comprehension of morpho-logical complexity. In L. Carlson, C. Hoelscher, and T. Shipley (eds), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 1643–1648.
- Kuperman, V. and Van Dyke, J. (2013). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language* 65(1): 42–73.
- Kuperman, V., Bertram, R., and Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes* 23: 1089–1132.
- Kuperman, V., Bertram, R., and Baayen, R. H. (2010). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language* 62: 83–97.
- Kuperman, V., Schreuder, R., Bertram, R., and Baayen, R. H. (2009). Reading of multimorphemic Dutch compounds: towards a multiple route model of lexical processing. *Journal of Experimental Psychology: HPP* 35: 876–895.
- Kuperman, V., Drieghe, D., Keuleers, E., and Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *Quarterly Journal of Experimental Psychology* 66(3): 563–580.
- Kuprijanova, Z. N., Barmiĉ, M. J. and Homiĉ, L. V. (1985), *Nenetskij jazyk [Nenets]*. Leningrad: Prosveščenie.
- Kurath, H., Lewis, R. E., and Kuhn, S. M. (eds) (2001). *Middle English Dictionary*. Ann Arbor: University of Michigan Press.
- Kuriaki, K. (2007). *A Grammar of Modern Indo-European*. European Union: Asociacion Cultural Dnghu.
- Kurusu, K. (2001). The phonology of morpheme realization. Ph.D. thesis. University of California at Santa Cruz, Santa Cruz, CA.
- Kuryłowicz, J. (1936). Dérivation lexicale et dérivation syntaxique. *Bulletin de la Société de Linguistique de Paris* 37: 79–92.
- Kuteva, T. (1998). Large linguistic areas in grammaticalization: auxiliation in Europe. *Language Sciences* 20(3): 289–311.

- Kuteva, T. (2000). Areal grammaticalization: The case of the Bantu-Nilotic borderland. *Folia Linguistica* 34(3-4): 267-83.
- Kutsch Lojenga, C. (1994). *Ngiti: A Central-Sudanic Language of Zaire*. Cologne: Rüdiger Köppe.
- Laakso, J. (1997). On verbalising nouns in Uralic. *Finnisch-Ugrische Forschungen* 54(3): 267-304.
- Laakso, J. (2004). Derivation, morphopragmatics, and language contact—on the role of German influence in Estonian word-formation. In I. Hyvärinen, Petri Kallio, and J. Korhonen (eds), *Etymologie, Entlehnungen und Entwicklungen. Festschrift für Jorma Koivulehto zum 70. Geburtstag*, Mémoires de la Société Néophilologique de Helsinki, LXIII. Helsinki: Société Néophilologique, 171-180.
- Laakso, J. (forthcoming). The history of word-formation in Uralic. In P. O. Müller, I. Ohnheiser, S. Olsen, and F. Rainer (eds), *Word-Formation: An International Handbook of the Languages of Europe*. Berlin: De Gruyter Mouton.
- Laalo, K. (1998). Diminutives in Finnish child-direct and child speech. In S. Gillis (ed.), *Studies in the Acquisition of Number and Diminutive Marking. Antwerp Papers in Linguistics* 95: 137-148.
- Laalo, K. (2001). Diminutives in Finnish child-directed and child speech: morphopragmatic and morphophonemic aspects. *Psychology of Language and Communication* 5(2): 71-80.
- Labelle, M. (1992). Change of state and valency. *Journal of Linguistics* 28: 375-414.
- Labov, W. (1973). The boundaries of words and their meanings. In C.-J. N. Bailey and R. W. Shuy (eds), *New Ways of Analyzing Variation in English*. Washington, DC: Georgetown University Press, 340-373.
- Laca, B. (1986). *Die Wortbildung als Grammatik des Wortschatzes. Untersuchungen zur spanischen Subjektnominalisierung*. Tübingen: Narr.
- Ladányi, M. (2007). *Produktivitás és analógia a szóképzésben: elvek és esetek* [Productivity and Analogy in Word-formation: Principles and Case Studies]. Budapest: Tinta Kiadó.
- Lafone Quevedo, S. A. (1895). Introducción y notas. Grupo Mataco-Mataguay del Chaco. Dialecto Noctén por Inocencio Massei. *Boletín del Instituto Geográfico Argentino*, XVI: 37-124.
- Lafone Quevedo, S. A. (1896). Introducción, notas y comentarios. Grupo Mataco-Mataguay del Chaco. Dialecto Vejoz. Con vocabulario y apuntes de D'Orbigny. *Boletín del Instituto Geográfico Argentino*, XVII, Cuadernos 4(5-6): 121-176.
- Lafone Quevedo, S. A. (1912). Pronominal classification of certain South American Indian stocks. Buenos Aires: Coni Hermanos.
- Lahaie, K. (1985). The ma-prefix in Afroasiatic. In Russell G. Schuh (ed.), *Précis from the 15th Conference on African Linguistics (University of California, Los Angeles, 1984)*. *Studies in African Linguistics*. Supplement 9: 186-190.
- Lahiri, A. and Marslen-Wilson, W. (1991). The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition* 38: 245-294.
- Lakoff, G. (1965). On the nature of syntactic irregularity. Ph.D. dissertation, Indiana University, Bloomington, IN. Published 1970 as *Irregularity in Syntax*. New York: Holt, Rinehart, and Winston.
- Lakoff, G. (1987). *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

- Lalou, M. (1953). Tibétain ancien Bod/Bon. *Journal Asiatique* 241: 275–276.
- Lang, M. F. (1990). *Spanish Word Formation: Productive Derivational Morphology in the Modern Lexis*. London: Routledge.
- Langacker, R. W. (1977). *Studies in Uto-Aztecan Grammar I: An Overview of Uto-Aztecan Grammar*. Dallas, TX: Summer Institute of Linguistics.
- Langacker, R. W. (1991). *Foundations of Cognitive Grammar: Descriptive Application*. Stanford: Stanford University Press.
- Langdon, M. (1970). *A Grammar of Diegueno*. Berkeley: University of California Press.
- LaPolla, R. J. (2003). Overview of Sino-Tibetan morphosyntax. In G. Thurgood and R. J. LaPolla (eds), *The Sino-Tibetan Languages*. London: Routledge, 22–42.
- LaPolla, R. J. (2012). Comments on methodology and evidence in Sino-Tibetan comparative linguistics. *Language and Linguistics* 13(1): 117–132.
- Lappe, S. (2003). Monosyllabicity in prosodic morphology: the case of truncated personal names in English. *Yearbook of Morphology, 2002*. Dordrecht: Kluwer Academic, 135–186.
- Lappe, S. (2007). *English Prosodic Morphology*. Dordrecht: Springer.
- Larson, R. K. (1985). Bare NP-adverbs. *Linguistic Inquiry* 16: 595–621.
- Larson, R. K. (1987). ‘Missing prepositions’ and the analysis of English free relative clauses. *Linguistic Inquiry* 18: 239–266.
- Lass, R. (1997). *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
- Laudanna, A. and Burani, C. (1995). Distributional properties of derivational affixes: implications for processing. In L. B. Feldman (ed.), *Morphological Aspects of Language Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates, 345–364.
- Laudanna, A., Burani, C., and Cermele, A. (1994). Prefixes as processing units. *Language and Cognitive Processes* 9: 295–316.
- Laufer, B. (1914). *Chinese Clay Figures*. Chicago: Field Museum of Natural History.
- Laughren, M., Hale, K., and Warlpiri Lexicography Group (2007). *Warlpiri-English Encyclopaedic Dictionary*. Brisbane: University of Queensland.
- Launey, M. (2002). On some causative doublets in Classical Nahuatl. In M. Shibatani (ed.), *The Grammar of Causation and Interpersonal Manipulation*. Amsterdam: Benjamins, 301–317.
- Lavrent'ev, J. I. (1984). *Sostav slova i slovoobrazovanie v sovremennom marijskom jazyke* [Compounding and derivation in contemporary Mari]. Joškar-Ola: Marijskoe knižnoe izdatel'stvo.
- Lavric, A., Clapp, A., and Rastle, K. (2007). ERP evidence of morphological analysis from orthography: A masked priming study. *Journal of Cognitive Neuroscience* 19(5): 866–877.
- Lavric, A., Elchlepp, H., and Rastle, K. (2012). Tracking hierarchical processing in morphological decomposition with brain potentials. *Journal of Experimental Psychology: Human Perception and Performance* 38(4): 811.
- Lazar, O. (1975). *The Formation of Abstract Nouns in the Uralic Languages*. Stockholm: Almqvist and Wiksell.
- Lazzeroni, R. (1998). Sanskrit. In A. Giacalone Ramat and P. Ramat (eds), *The Indo-European Languages*. London: Routledge, 98–124.
- Le Bleis, Y. and Barreteau, D. (1987). Les extensions verbales en Mafa. In H. Jungraithmayr and H. Tourneux (eds), *Etudes tchadiques. Classes et extensions verbales*. Paris: Paul Geuthner, 99–114.

- Lee, J. (ed.) (2004). *Kriol-Inglish Dikshenri: Kriol-English Dictionary*. Draft: October 2004 (ASEDA item 0739 Kriol Dictionary, Canberra).
- Leer, J. (1979). *Proto-Athabaskan Verb Stem Variation: I. Phonology*. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks.
- Lees, R. B. (1960). *The Grammar of English Nominalizations*. Bloomington: Indiana University Press and Den Haag: Mouton.
- Lehman, C. and Moravcsik, E. (2000). Noun. In *Morphology: An International Handbook on Inflection and Word Formation*. Berlin: Mouton de Gruyter, 733–757.
- Lehmann, C. (2005). Pleonasm and hypercharacterisation. *Yearbook of Morphology* 2005. Dordrecht: Kluwer Academic, 119–154.
- Lehnert, M. (1971). *Reverse Dictionary of Present-Day English*. Leipzig: VEP Verlag Enzyklopädie.
- Lehrer, A. (2003). Polysemy in derivational affixes. In B. Nerlich, Z. Todd, V. Herman, and D. D. Clarke (eds), *Polysemy. Flexible Patterns of meaning in Mind and Language*. Berlin: Mouton de Gruyter, 217–232.
- Lehtisalo, T. (1936). *Über die primären uralischen Ableitungssuffixe*. Helsinki: Finno-Ugrian Society.
- Leisi, E. (1985). *Praxis der englischen Grammatik*. 2nd edn. Heidelberg: Winter.
- Leitner, G. (1974). *Denominale Verbalisierungen im Englischen*. Tübingen: Niemeyer.
- Lemhoefer, K., Dijkstra, A., Schriefers, H., Baayen, R., Grainger, J., and Zwitserlood, P. (2008). Native language influences on word recognition in a second language: a megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34: 12–31.
- Lemhöfer, K., Koester, D., and Schreuder, R. (2011). When bicycle pump is harder to read than bicycle bell: effects of parsing cues in first and second language compound reading. *Psychonomic Bulletin and Review* 18(2): 364–370.
- Leopold, W. F. (1949). *Speech Development of a Bilingual Child* (4 vols). Evanston, IL: Northwestern University Press.
- Lepämaa, A.-L., Lieko, A., and Silfverberg, L. (1996). *Miten sanoja johdetaan: Suomen kielen johto-oppia* [How to derive words: Derivation in the Finnish language]. Helsinki: Finn Lectura.
- Leslau, W. (1945). The influence of Cushitic on the Semitic languages of Ethiopia: a problem of substratum. *Word* 1: 59–82.
- LeSourd, P. S. (1995). Diminutive verb forms in Passamaquoddy. *International Journal of American Linguistics* 61(1): 103–134.
- Levelt, W. J. M., Roelofs, A. and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22: 1–38.
- Levi, J. N. (1978). *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.
- Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago, IL: University of Chicago Press.
- Levin, B. and Rappaport, M. (1986). The formation of adjectival passives. *Linguistic Inquiry* 17: 623–661.
- Levin, B. and Rappaport Hovav, M. (1991). Wiping the slate clean: a lexical semantic exploration. In B. Levin and S. Pinker (eds), *Lexical and Conceptual Semantics*. Oxford: Blackwell, 123–151.
- Levin, B. and Rappaport Hovav, M. (1995). *Unaccusativity: At the Syntax–Lexical Semantics Interface*. Cambridge, MA: MIT Press.
- Levin, B. and Rappaport Hovav, M. (1998). Morphology and lexical semantics. In A. Zwicky and A. Spencer (eds), *Handbook of Morphology*. Oxford: Blackwell, 248–271.

- Levin, B. and Rappaport Hovav, M. (2003). Roots and templates in the representation of verb meaning. Handout of talk at the Department of Linguistics, Stanford University, May 15.
- Levinson, L. (2007). The roots of verbs. Ph.D. dissertation. NYU, New York.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition* 106: 1126–1177.
- Lewandowska-Tomaszczyk, B. (2007). Polysemy, prototypes, and radial categories. In D. Geeraerts and H. Cuyckens (eds), *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press, 139–169.
- Lewis, D. J. and Windsor, J. (1996). Children's analysis of derivational suffix meanings. *Journal of Speech and Hearing Research* 39: 209–216.
- Lewis, G. L. (1967). *Turkish Grammar*. Oxford: Clarendon Press.
- Lewis, G., Solomyak, O., and Marantz, A. (2011). The neural basis of obligatory decomposition of suffixed words. *Brain and language* 118(3): 118.
- Lewis, M. P. (ed.) (2009). *Ethnologue: Languages of the World*. 16th edn. Dallas, TX: SIL International. Online version: <<http://www.ethnologue.com/>>.
- Lewis, M. P., Simons, G. F., and Fleming, C. D. (eds) (2013). *Ethnologue: Languages of the World*. 17th edn. Dallas, TX: SIL International. <<http://www.ethnologue.com>>.
- Li, J. (1996). Bugar: a new Mon-Khmer language of Yunnan Province, China. *Mon-Khmer Studies* 26: 135–159.
- Li, Y. (李云兵). (2005). *Bugeng yu yan jiu* (布赉语研究) (A Study of Bugar). Beijing: Min zu chu ban she (民族出版社).
- Libben, G. (1994). How is morphological decomposition achieved? *Language and Cognitive Processes* 9(3): 369–391.
- Libben, G. (2010). Compound words, semantic transparency and morphological transcendence. In S. Olsen (ed.), *New Impulses in Word-Formation*. Hamburg: Buske, 317–330.
- Libben, G., Derwing, B. and de Almeida, R. (1999). Ambiguous novel compounds and models of morphological parsing. *Brain and Language* 68: 378–386.
- Libben, G., Buchanan, L. and Colangelo, A. (2003). Morphology, semantics and aphasia: the failure of inhibition hypothesis. *Logos and Language* 4(1): 45–53.
- Lichtenberk, F. (1983). *A Grammar of Manam*. Oceanic Linguistics Special Publication 18. Honolulu: University of Hawaii Press.
- Lieber, R. (1980). On the organization of the lexicon. Ph.D. dissertation. MIT.
- Lieber, R. (1981). Morphological conversion within a restricted theory of the lexicon. In M. Moortgat, H. van der Hulst, and T. Hoekstra (eds), *The Scope of Lexical Rules*. Dordrecht: Foris, 161–200.
- Lieber, R. (1983). Argument linking and compounds in English. *Linguistic Inquiry* 14: 251–286.
- Lieber, R. (1984). Consonant gradation in Fula: an autosegmental approach. In M. Aronoff and R. Oehrle (eds), *Language Sound Structure*. Cambridge, MA: MIT Press, 329–345.
- Lieber, R. (1987). *An Integrated Theory of Autosegmental Processes*. Albany: SUNY Press.
- Lieber, R. (1992). *Deconstructing Morphology*. Chicago: University of Chicago Press.
- Lieber, R. (1998). The suffix *-ize* in English: implications for morphology. In S. G. Lapointe, D. K. Brentari, and P. M. Farrell (eds), *Morphology and Its Relation to Phonology and Syntax*. Stanford, CA: CSLI Publications, 12–34.
- Lieber, R. (2004). *Morphology and Lexical Semantics*. Cambridge: Cambridge University Press.
- Lieber, R. (2005). Word formation processes in English. In P. Štekauer and R. Lieber (eds), *Handbook of Word Formation*. Dordrecht: Springer, 375–422.

- Lieber, R. (2006). The category of roots and the roots of categories: what we learn from selection in derivation. *Morphology* 16: 247–272.
- Lieber, R. (2009a). A lexical semantic approach to compounding. In R. Lieber and P. Štekauer (eds), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 78–104.
- Lieber, R. (2009b). Towards an OT morphosemantics: the case of *-hood*, *-dom*, and *-ship*. In S. Olsen (ed.), *New Impulses in Word-formation*. Hamburg: Buske, 61–80.
- Lieber, R. (2010a). *Introducing Morphology*. Cambridge: Cambridge University Press.
- Lieber, R. (2010b). On the lexical semantics of compounds: non-affixal (de)verbal compounds. In S. Scalise and I. Vogel (eds), *Cross-Disciplinary Issues in Compounding*. Amsterdam: Benjamins, 127–144.
- Lieber, R. and Štekauer, P. (eds) (2009). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press.
- Lin, Xiangrong (1993). 嘉戎語研究 *Jiarongyu Yanjiu* [A Grammar of Rgyalrong]. Chengdu: Sichuan Nationalities Press.
- Lin, Y.-H. (1989). Autosegmental treatment of segmental processes in Chinese phonology. Ph.D. dissertation. University of Texas at Austin.
- Lin, Y.-H. (2004). Chinese affixal phonology: some analytical and theoretical issues. *Language and Linguistics* 5(4): 1019–1046.
- Lin, Y.-H. (2008). The morphophonology of *Pingding* Er infixation: an OT analysis. In Y. E. Hsiao, H.-C. Hsu, L.-H. Wee, and D.-A. Ho (eds), *Interfaces in Chinese Phonology: Festschrift in Honor of Matthew Y. Chen on his 70th Birthday*. Taipei: Institute of Linguistics, Academia Sinica, 161–182.
- Lin, You-Jing (2010). By no means marginal: privative tone in Zhuokeji Rgyalrong. *Language and Linguistics* 13(4): 625–662.
- Lin, Y.-J. and Luo, E. (2003). 茶堡嘉戎語大藏話的趨向前綴與動詞詞幹的變化 Chabao Jiarongyu Dazanghua de quxiang qianzhui yu dongci cigan de bianhua [Orientation prefixes and verb stem variation in the Dazang variety of Chabao rGyalrong]. *Minzu Yuwen* 4: 19–29.
- Lindsay, M. and Aronoff, M. (2013). Natural selection in self-organizing morphological systems. In G. Boyé, F. Montermini, and J. Tseng (eds), *Selected Proceedings of the 7th Décembrettes*. Munich: Lincom Europa, 133–153.
- Lionnet, A. (2001). *Los elementos del Tarahumar y otros Estudios Lingüísticos*, Segundo Volumen. Sisogichi, Mexico: Ediciones Diocesanas de la Tarahumara.
- Lipiński, E. (2001). *Semitic Languages: Outline of a Comparative Grammar* (Orientalia Lovaniensia Analecta 80). Leuven: Peeters.
- Lipka, L. (1971). Grammatical categories, lexical items and word-formation. *Foundations of Language* 7: 211–238.
- Lipka, L. (1990). *An Outline of English Lexicology*. Tübingen: Niemeyer.
- Lipka, L., Handl, S., and Falkner, W. (2004). Lexicalization and institutionalization. The state of the art in 2004. *SKASE Journal of Theoretical Linguistics* 1(1): 2–19.
- Lloyd, C. (2005). Some Latinate deverbal suffixes in Middle English—their integration, productivity and semantic coherence <http://etheses.whiterose.ac.uk/689/1/uk_bl_ethos_422633.pdf>, accessed February 21, 2013.
- Lloyd, C. (2006). Experience or experiment? some distinctions between French nominal suffixes in Middle English. In N. Ritt et al. (eds), *Medieval English and Its Heritage: Structure, Meaning and Mechanisms of Change* (Studies in English Medieval Language and Literature: 16). Frankfurt: Peter Lang, 179–195.

- Lloyd, C. (2007a). From denominal to deverbal: the suffix *-age* in Middle English. In G. Mazzon (ed.), *Studies in Middle English Forms and Meanings* (Studies in English Medieval Language and Literature: 19). Frankfurt: Peter Lang, 145–164.
- Lloyd, C. (2007b). From Middle English to Shakespeare: semantics of a French deverbal suffix in English. *English Studies: A Journal of English Language and Literature* 88(5): 601–620.
- Lo Duca, M. G. (1990). *Creatività et regole: studio sull'acquisizione della morfologia derivativa dell'italiano* [Creativity and rules: study of the acquisition of derivational morphology in Italian]. Bologna: Il Mulino.
- Lombardi, L. and McCarthy, J. (1991). Prosodic circumscription in Choctaw morphology. *Phonology* 8: 37–72.
- Loprieno, A. (1995). *Ancient Egyptian: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Loprieno, A. and Müller, M. (2012). Ancient Egyptian and Coptic. In Z. Frajzyngier and E. Shay (eds), *The Afroasiatic Languages*. Cambridge: Cambridge University Press, 102–144.
- Loukotka, Č. (1968). *Classification of South American Indian Languages*. Los Angeles: University of California.
- Louw, J. A. (1975). Palatalization of bilabials in the passive, diminutive and locative in Xhosa and Tsonga. *Afrika und Übersee* 59: 241–278.
- Lovick, O. and Rice, K. (forthcoming). Slavey (Dene) and other Athabaskan languages. In N. Grandi and L. Körtvélyessy (eds), *Edinburgh Handbook of Evaluative Morphology*. Edinburgh: Edinburgh University Press.
- Lowe, R. (1983). *Kangiryuarmit Uqauhingita Numiktittitdjutingit: Basic Kangiryuarmit Eskimo Dictionary*. Inuvik: Committee for Original Peoples Entitlement.
- Lowe, R. (1984). *Uummarmiut Uqalaungiha Mumikhitchi utingit: Basic Uummarmiut Eskimo Dictionary*. Inuvik: Committee for Original Peoples Entitlement.
- Lüdtke, J. (2005). *Romanische Wortbildung. Inhaltlich—diachronisch—synchronisch*. Tübingen: Stauffenburg.
- Lukas, J. (1937a). *A Study of the Kanuri Language: Grammar and Vocabulary*. London: Oxford University Press.
- Lukas, J. (1937b). *Zentralsudanische Studien: Wörterverzeichnisse der Deutschen Zentral-Afrika Expedition 1910/11*. Hamburg: Friederichsen de Gruyter.
- Luschützky, H. C. and Rainer, F. (2011). Agent noun polysemy in a cross-linguistic perspective. *Language Typology and Universals* 64(4): 287–338.
- Luschützky, H. C. and Rainer, F. (2013). Instrument and place nouns: a typological and diachronic perspective. *Linguistics* 51(6): 1301–1359.
- Lustig, A. (2010). *A Grammar and Dictionary of Zaiwa*. Leiden: Brill.
- Lüttmann, H., Zwisserlood, P., Böhl, A., and Bölte, J. (2011). Evidence for morphological composition at the form level in speech production. *Journal of Cognitive Psychology* 23(7): 818–836.
- Lutz, A. (1997). Sound change, word formation and the lexicon: the history of the English prefix verbs. *English Studies: A Journal of English Language and Literature* 78(3): 258–290.
- Lutz, A. (2002). When did English begin? In T. Fanego, B. Méndez-Naya, and E. Seoane (eds), *Sounds, Words, Texts and Change* (Amsterdam Studies in the Theory and History of Linguistic Science IV: Current Issues in Linguistic Theory: 224). Amsterdam: Benjamins, 145–71.

- Lydall, J. (1988). Gender, number and size in Hamar. In M. Bechhaus-Gerst and F. Serzisko (eds), *Cushitic and Omotic. Papers from the International Symposium on Cushitic and Omotic Languages, Cologne, January 6–9, 1986*. Hamburg: Helmut Buske, 77–90.
- Lynch, J. and Horoi, R. (2002). Arosi. In J. Lynch, M. Ross, and T. Crowley (eds), *The Oceanic Languages*. Richmond: Curzon Press, 562–572.
- Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.
- McCarthy, J. J. (1979). Formal problems in Semitic phonology and morphology. Ph.D. thesis. MIT, Cambridge, MA.
- McCarthy, J. J. (1981). A prosodic theory of non-concatenative morphology. *Linguistic Inquiry* 12: 373–418.
- McCarthy, J. J. (1982). Prosodic structure and expletive infixation. *Language* 58: 574–590.
- McCarthy, J. J. (1983). Consonantal morphology in the Chaha verb. *Proceedings of the West Coast Conference on Formal Linguistics* 2: 176–188.
- McCarthy, J. J. (1984). Prosodic organization in morphology. In M. Aronoff and R. Oehrle (eds), *Language Sound Structure*. Cambridge, MA: MIT Press, 299–317.
- McCarthy, J. J. (1986). Lexical phonology and nonconcatenative morphology in the history of Chaha. *Revue québécoise de linguistique* 16: 209–228.
- McCarthy, J. J. (1993). Templatic form in prosodic morphology. *Proceedings of the Formal Linguistics Society of Mid-America* 3: 187–218.
- McCarthy, J. J. (2000). The prosody of phase in Rotuman. *Natural Language and Linguistic Theory* 18: 147–197.
- McCarthy, J. J. (2008). *Doing Optimality Theory*. Malden, MA: Wiley-Blackwell.
- McCarthy, J. J. and Prince, A. (1986). Prosodic morphology. MS. University of Massachusetts and Brandeis University.
- McCarthy, J. J. and Prince, A. (1990). Foot and word in prosodic morphology: the Arabic broken plural. *Natural Language and Linguistic Theory* 8: 209–282.
- McCarthy, J. J. and Prince, A. (1993a). Generalized alignment. In *Yearbook of Morphology, 1993*. Dordrecht: Kluwer Academic, 79–153.
- McCarthy, J. J. and Prince, A. (1993b). Prosodic morphology I: constraint interaction and satisfaction. MS. University of Massachusetts, Amherst and Rutgers University.
- McCarthy, J. J. and Prince, A. (1994a). The emergence of the unmarked. In M. Gonzalez (ed.), *Proceedings of NELS 24*. Amherst, MA: GLSA, 333–379.
- McCarthy, J. J. and Prince, A. (1994b). An overview of prosodic morphology. Part I: Templatic form in reduplication. Part II: Template satisfaction. Paper presented at Workshop on Prosodic Morphology, Utrecht University.
- McCarthy, J. J. and Prince, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Dickey, and S. Urbanczyk (eds), *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. Amherst, MA: GLSA, 249–384.
- McCarthy, J. J. and Prince, A. (1996). *Prosodic Morphology 1986*. Technical Report #32. Rutgers University Center for Cognitive Science.
- McCawley, J. D. (1968). Lexical insertion in a transformational grammar without deep structure. In *Proceedings of the Fourth Regional Meeting of the Chicago Linguistic Society*. Chicago, IL: Chicago Linguistic Society, 71–80.
- McCawley, J. D. (1988). *The Syntactic Phenomena of English*. Chicago: University of Chicago Press.
- McClelland, J. L. and Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology* 18: 1–86.

- McClelland, J. L. and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of the basic findings. *Psychological Review* 88: 375–407.
- McConchie, R. W. (2000). Fashionable idiolects? The use of the negative prefix *dis-* 1520–1620. In D. Kastovsky and A. Mettinger (eds), *The History of English in a Social Context: A Contribution to Historical Sociolinguistics* (Trends in Linguistics: Studies and Monographs: 129). Berlin: Mouton de Gruyter, 199–225.
- McConchie, R. W. (2007). Dis-history: modelling the introduction and diffusion of a borrowed prefix in Middle English. In M. Rissanen et al. (eds), *Change in Meaning and the Meaning of Change: Studies in Semantics and Grammar from Old to Present-Day English* (Mémoires de la Société Néophilologique de Helsinki (MSNH): 72). Helsinki: Société Néophilologique, 345–372.
- Macdonald, R. R. and Soenjono Dardjowidjojo. (1967). *A Student's Reference Grammar of Modern Formal Indonesian*. Washington, DC: Georgetown University Press.
- Macdonell, A. A. (1916). *A Vedic Grammar for Students*. Oxford: Oxford University Press.
- McGinnis, M. (2008). Applicatives. *Language and Linguistics Compass* 2: 1225–1245.
- McGregor, W. B. (2000). *Cockatoos, Chaining-horsemen, and Mud-eaters*. *Anthropos* 95: 3–22. <<http://www.jstor.org/stable/40465858>> accessed July 15, 2012.
- McGregor, W. (2002). *Verb Classification in Australian Languages*. Berlin: Mouton de Gruyter.
- McKoon, G., and Macfarland, T. (2000). Externally and internally caused change of state verbs. *Language* 76: 833–858.
- McLaughlin, J. E. (1987). A phonology and morphology of Panamint. Ph.D. dissertation. University of Kansas.
- McMahon, A. M. S. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Meakins, F., McConvell, P., Charola, E., McNair, N., McNair, H., and Campbell, L. (2013). *Gurindji to English Dictionary*. Batchelor, NT: Batchelor Press..
- McPherson, L. and Paster, M. (2009). Evidence for the mirror principle and morphological templates in Luganda affix ordering. In A. Ojo and L. Moshi (eds), *Selected Proceedings of the 39th Annual Conference on African Linguistics*. Somerville, Massachusetts: Cascadilla Proceedings Project, 56–66.
- McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39: 21–46.
- McQueen, J. M. and Cutler, A. (1998). Morphology in word recognition. In A. Spencer (ed.), *Handbook of Morphology*. Oxford: Blackwell, 406–427.
- MacWhinney, B. (1978). The acquisition of morphophonology. *Monograph of the Society for Research in Child Development*, 43 (Serial No. 174).
- Magnusson, R. (1954). *Studies in the Theory of the Parts of Speech* (Lund Studies in English XXIV). Lund: C.W.K. Gleerup.
- Mahlow, C. and Piotrowski, M. (eds) (2009). *State of the Art in Computational Morphology*. Berlin: Springer.
- Maiden, M. and Robustelli, C. (2000). *A Reference Grammar of Modern Italian*. London: Arnold.
- Majtinskaja, K. E. (1974). Sravnitel'naja morfologija finno-ugorskix jazykov [Comparative morphology of Finno-Ugric languages]. In *Osnovy finno-ugorskogo jazykoznanija: Voprosy proisxoždenija i razvitija finno-ugorskix jazykov*. Moskva: Nauka, 214–382.
- Malchukov, A. (1995). *Even*. München; Newcastle: Lincom Europa.

- Malik, A. N. (1995). *The Phonology and Morphology of Panjabi*. New Delhi: Munshiram Manoharlal Publishers Pvt Ltd.
- Mallory, J. P. and Adams, D. Q. (2006). *Oxford Introduction to Proto- Indo-European and the Proto-Indo-European World*. Oxford: Oxford University Press.
- Man, E. H. (1889). *A Dictionary of the Central Nicobarese Language (English-Nicobarese and Nicobarese-English)*. London: W.H. Allen and Co.
- Manaster Ramer, A. (1992). A Northern Uto-Aztecan sound law: *-c- > -y-. *International Journal of American Linguistics* 58: 251–268.
- Manley, T. M. (1972). *Outline of Sre Structure*. Oceanic Linguistics Special Publication No. 12.
- Manning, C. (1996). *Ergativity: Argument Structure and Grammatical Relations*. Stanford, CA: CSLI Publications.
- Manova, S. (2010). Suffix combinations in Bulgarian: parsability and hierarchy-based ordering. *Morphology* 20: 267–296.
- Manova, S. and Aronoff, M. (2010). Modeling affix order. *Morphology* 20: 109–131.
- Manova, S. and Dressler, W. U. (2005). The morphological technique of conversion in the inflecting-fusional type. In L. Bauer and S. Valera (eds), *Approaches to Conversion/Zero-derivation*. Münster: Waxmann, 67–102.
- Marantz, A. (1982). Re reduplication. *Linguistic Inquiry* 13(3): 435–482.
- Marantz, A. (1997). No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams (eds), *University of Pennsylvania Working Papers in Linguistics* 4(2): 201–225.
- Marantz, A. (1999). Reconstructing the lexical domain with a single generative engine. MS. MIT.
- Marantz, A. (2001). *Words*. MS. MIT.
- Marchand, H. (1960/1969). *The Categories and Types of Present-Day English Word Formation. A Synchronic-Diachronic Approach*. Munich: Beck.
- Marchand, H. (1963a). On a question of contrary analysis with derivationally connected but morphologically uncharacterized words. *English Studies* 44: 176–187.
- Marchand, H. (1963b). On content as a criterion of derivational relationship with backderived words. *Indogermanische Forschungen* 68: 170–175.
- Marchand, H. (1964). A set of criteria for the establishing of derivational relationship between words unmarked by derivational morphemes. *Indogermanische Forschungen* 69: 10–19.
- Marchand, H. (1966). On attributive and predicative derived adjectives and some problems related to the distinction. *Anglia* 84(2): 131–149.
- Markianova, L. F. (1985). *Glagol'noe slovoobrazovanie v karel'skom jazyke* [Verbal derivation in Karelian]. Petrozavodsk: s.n.
- Marlow, P. (2000). Koyukon sound system and orthography. In J. Jetté and E. Jones (J. Kari, editor-in-chief), *Koyukon Athabaskan Dictionary*. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks, lxvi–lxxi.
- Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition. In H. Bouma and D. G. Bouwhuis (eds), *Attention and Performance: Control of Language Processes*, vol. X. Hillsdale, NJ: Lawrence Erlbaum Associates, 125–150.
- Marslen-Wilson, W. D. (1996). Function and process in spoken word recognition. In H. Bouma and D. G. Bouwhuis (eds), *Attention and Performance: Control of Language Processes*, vol. X. Hillsdale, NJ: Lawrence Erlbaum Associates, 125–50.
- Martin, J. (1988). Subtractive morphology as dissociation. *Proceedings of the West Coast Conference on Formal Linguistics* 7: 229–240.

- Martin, J. B. and Munro, P. (2005). Proto-Muskogean morphology. In H. K. Hardy and J. Scancarelli (eds), *Native Languages of the Southeastern United States*. Lincoln, NE: University of Nebraska Press, 299–320.
- Martin, S. E. (1961). *Dagur Mongolian Grammar, Texts, and Lexicon* (Indiana University Publications, Uralic and Altaic series 4). Indiana University, Bloomington. The Hague: Mouton.
- Martin, S. E. (2004). *A Reference Grammar of Japanese*. Honolulu: University of Hawai'i Press.
- Martin Arista, J. (2006). Alternations, relatedness and motivation: Old English A-. In P. Guerrero Medina and E. Martínez Jurado (eds), *Where Grammar Meets Discourse: Functional and Cognitive Perspectives*. Córdoba: Universidad de Córdoba, 113–132.
- Masica, C. P. (1991). *The Indo-Aryan Languages*. New York: Cambridge University Press.
- Matisoff, J. A. (1972). Lahu nominalization, relativization, and genitivization. In J. P. Kimball (ed.), *Syntax and Semantics*. Vol. 1. New York: Academic Press, 237–257.
- Matisoff, J. A. (1973). Tonogenesis in Southeast Asia. In L. M. Hyman (ed.), *Consonant Types and Tone* (Southern California Occasional papers in Linguistics, No. 1). Los Angeles: UCLA, 71–95.
- Matisoff, J. A. (1991). The mother of all morphemes. augmentatives and diminutives in areal and universal perspective. *Papers from the First Annual Meeting of the Southeast Asian Linguistic Society, Tempe*, 293–349.
- Matisoff, J. A. (2003). Aslian: Mon-Khmer of the Malay Peninsula. *Mon-Khmer Studies* 33: 1–58.
- Matras, Y. and Sakel, J. (2007). Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31(4): 829–865.
- Mattes, V. (2006). One form—opposite meanings? Attenuative and intensive interpretation of full reduplication in Bikol, oral presentation, International Conference on Austronesian Linguistics 10, 2006, Puerto Princesa, The Philippines. Paper available online at <http://reduplication.uni-graz.at/texte/paper_ICAL.pdf>.
- Matthews, P. H. (1974). *Morphology. An Introduction to the Theory of Word Structure*. Cambridge: Cambridge University Press.
- Mayerthaler, W. (1977). *Studien zur theoretischen und zur französischen Morphologie*. Tübingen: Niemeyer.
- Means, N. (1998). *Temiar-English English-Temiar dictionary*. St. Paul, MN: Hamline University Press.
- Means, N. and Means, P. B. (1986). *Sengoi-English English-Sengoi Dictionary*. The Joint Centre on Modern East Asia, University of Toronto.
- Mediano, Z. D. (1976). Preliminary studies in the acquisition of Portuguese morphology by Brazilian children. Ph.D. dissertation. University of New Mexico, Albuquerque, NM.
- Megerdooomian, K. (2002). Beyond words and phrases: a unified theory of predicate composition. Ph.D. dissertation. University of Southern California, Los Angeles, CA.
- Mei, T.-L. 梅祖麟 (1980). 四聲別義中的時間層次 Si sheng bie yi zhong de shijian cengci [chronological strata in derivation by tone change]. *中國語文 Zhongguo Yuwen* 6: 427–443.
- Mei, T.-L. 梅祖麟 (2012). The causative *s- and nominalizing *-s in Old Chinese and related matters in Proto-Sino-Tibetan. *Language and Linguistics* 13(1): 1–28.
- Meillet, A. (1925). *La méthode comparative en linguistique historique*. Oslo: H. Aschehoug and Co.
- Mel'cuk, I. A. (1993). *Cours de morphologie générale*. Montréal: Les Presses de l'Université de Montréal.
- Melis, A. (1999). Description du Masa (Tchad): Phonologie, syntaxe et dictionnaire encyclopédique. Ph.D. thesis. Université de Tours.

- Melissaropoulou, D. and Ralli, A. (2008). Headedness in diminutive formation: evidence from Modern Greek and its dialectal variation. *Acta Linguistica Hungarica* 55: 183–204.
- Melloni, C. (2011). *Event and Result Nominals: A Morpho-Semantic Approach*. Bern etc.: Lang.
- Melzi, G. and King, K. A. (2003). Spanish diminutives in mother-child conversations. *Journal of Child Language* 30: 281–304.
- Mendikoetxea, A. (1999). Construcciones inacusativas y pasivas. In I. Bosque and V. Demonte (eds), *Gramática descriptiva de la lengua Española*. Madrid: Editorial Espasa, 1575–1629.
- Menges, K. (1978). Problems of Tungus Linguistics. *Anthropos* 73: 367–400.
- Méresse-Polaert, J. (1969). *Etude sur le langage des enfants de six ans*. Neuchâtel: Delachaux and Niestlé.
- Merlan, F. (1982). *Mangarayi*. Amsterdam: North-Holland.
- Messineo, C. (2000). Estudio del toba hablado en la provincia del Chaco (Argentina), aspectos gramaticales y discursivos. Ph.D. dissertation. University of Buenos Aires.
- Messineo, C. and Cúneo, P. (2010). Modos de clasificación nominal en toba (guaycurú) y maká (mataguaya): zoonimia y fitonimia. In C. Messineo, G. Scarpa, and F. Tola (eds), *Léxico y categorización etnobiológica en grupos indígenas del Gran Chaco*. Santa Rosa: University of La Pampa, 27–64.
- Messineo, C. and Tacconi, T. (2010). Recursos de formación del léxico en maká (mataguayo): zoonimia y fitonimia. In C. Messineo, G. Scarpa, and F. Tola (eds), *Léxico y categorización etnobiológica en grupos indígenas del Gran Chaco*. Santa Rosa: University of La Pampa, 83–116.
- Meunier, F. and Segui, J. (1999a). Frequency effects in auditory word recognition: the case of suffixed words. *Journal of Memory and Language* 41: 327–344.
- Meunier, F. and Segui, J. (1999b). Morphological priming effect: the role of surface frequency. *Brain and Language* 68(1): 54–60.
- Meyer, H. A. E. (1843). *Vocabulary of the language spoken by the aborigines of the southern and eastern portions of the settled districts of South Australia, ..., by the tribes in the vicinity of Encounter Bay, and, with slight variations, by those extending along the coast to the eastward around Lake Alexandrina and for some distance up ... the River Murray, preceded by a grammar, showing the construction of the language as far as at present known*. Adelaide: James Allen.
- Milin, P., Filipović Durdević, D., and Moscoso del Prado Martín, F. (2009a). The simultaneous effects of inflectional paradigms and classes on lexical recognition: evidence from Serbian. *Journal of Memory and Language* 60(1): 50–64.
- Milin, P., Kuperman, V., Kostić, A., and Baayen, R. (2009b). Paradigms bit by bit: an information-theoretic approach to the processing of paradigmatic structure in inflection and derivation. In J. P. Blevins and J. Blevins (eds), *Analogy in Grammar: Form and Acquisition*. Oxford: Oxford University Press, 214–252.
- Miller, A. (2001). *A Grammar of Jamul Tiipay*. New York and Berlin: Mouton de Gruyter.
- Miller, D. G. (1997). The Morphological Legacy of French: Borrowed Suffixes on Native Bases in Middle English. *Diachronica: International Journal for Historical Linguistics/Revue Internationale pour la Linguistique Historique/Internationale Zeitschrift für Historische Linguistik* 14(2): 233–264.
- Miller, G. (2006). *Latin Suffixal Derivatives in English and Their Indo-European Ancestry*. Oxford: Oxford University Press.
- Miller, G. A. and Fellbaum, C. (1991). Semantic networks of English. In B. Levin and S. Pinker (eds), *Lexical and Conceptual Semantics*. Oxford: Blackwell, 197–229.

- Miller, K. (2008). Hadza Grammar Notes. 3rd International Symposium on Khoisan Languages and Linguistics, Riezlern.
- Miller, R. A. (1967). *The Japanese language*. Chicago: University of Chicago Press.
- Miller, R. A. (1994). The original geographic distribution of the Tungus languages. In H. I. Aronson (ed.), *Non-Slavic languages of the USSR. Papers from the Fourth Conference*. Columbus, OH: Slavica Publishers, Inc., 272–297.
- Miller, W. R. (1983). Uto-Aztecan languages. In A. Ortiz (ed.), *Handbook of North American Indians*, vol. 10—Southwest, Washington, DC: Smithsonian Institution, 113–124.
- Miller, W. R. (1984). The classification of the Uto-Aztecan languages based on lexical evidence. *International Journal of American Linguistics* 50: 1–24.
- Miller, W. R. (1996). *Guarijío: Gramática, textos, y vocabulario*. Mexico, DF: Universidad Nacional Autónoma de México.
- Milne, L. (1921). *An Elementary Palaung Grammar*. London: Oxford University Press.
- Minkova, D. (1991). *The History of Final Vowels in English: The Sound of Muting*. Berlin: Mouton de Gruyter.
- Minkova, D. (2008). Prefixation and stress in Old English: in memoriam Richard Hogg (1944–2007). *Word Structure* 1(1): 21–52.
- Mintz, M. W. (1971). *Bikol Grammar Notes*. PALI Language Texts. Honolulu: University of Hawaii Press.
- Mishra, H. (1975). *Historical Oriya Morphology*. Varanasi: Bharata Manisha.
- Mithun, M. (1984). On the evolution of noun incorporation. *Language* 60: 847–94.
- Mithun, M. (1986). On the nature of noun incorporation. *Language* 62(1): 32–37.
- Mithun, M. (1989). The acquisition of polysynthesis. *Journal of Child Language* 16: 285–312.
- Mithun, M. (1998). The sequencing of grammaticization effects. In M. Schmidt, J. Austin, and D. Stein (eds), *Historical Linguistics 1997*. Amsterdam: Benjamins, 291–314.
- Mithun, M. (1999). *The Languages of Native North America*. Cambridge: Cambridge University Press.
- Mithun, M. (2000). Incorporation. In G. Booij, C. Lehmann, and J. Mugdan (eds), *Morphology. An International Handbook on Inflection and Word-Formation*. Berlin: Walter de Gruyter, 916–928.
- Miwa, K., Libben, G., Dijkstra, T., and Baayen, R. H. (2014). The time-course of lexical activation in Japanese morphographic word recognition: evidence for a character-driven processing model. *Quarterly Journal of Experimental Psychology* 67: 79–113.
- Mohan, S. (2006). Echo-word formation in Hindi. *Indian Linguistics* 67: 119–26.
- Mohanan, K. P. (1986). *The Theory of Lexical Phonology*. Dordrecht: Reidel.
- Mohanan, T. (1995). Wordhood and lexicality: noun incorporation in Hindi. *Natural Language and Linguistics* 13: 75–134.
- Mohanan, T. and Mohanan, K. P. (1999). On representations in grammatical semantics. In T. Mohanan and L. Wee (eds), *Grammatical Semantics: Evidence for Structure in Meaning*. Stanford: CSLI Publications, 23–76.
- Möhlig, R. (2001). Zur Morphologie der altenglischen -nes-Ableitungen. Eine synchronische Wortbildungsanalyse. In T. Honegger, H. L. C. Tristram, and A. Fischer (eds), *Authors, Heroes and Lovers: Essays on Medieval English Literature and Language/Liebhaber, Helden und Autoren: Studien zur alt- und mittelenglischen Literatur und Sprache*. Collection Variations: 2. Bern: Peter Lang, 199–250.
- Mondorf, B. (2009). *More Support for More-support: The Role of Processing Constraints on the Choice between Synthetic and Analytic Comparative Forms*. Amsterdam: Benjamins.

- Montes Giraldo, J. J. (1976). El sistema, la norma y el aprendizaje de la lengua. *Thesaurus: Boletín del Instituto Caro y Cuervo* 31: 14–40.
- Moore, P. J. (2002). Point of view in Kaska historical narratives. Ph.D. dissertation. Indiana University.
- Moravcsik, E. (1978). Reduplicative constructions. In J. H. Greenberg (ed.), *Universals of Human Language*, vol. 3: *Word Structure*. Stanford, CA: Stanford University Press, 297–334.
- Moravcsik, E. (2000). Inflection. In G. Booij, C. Lehmann, and J. Mugdan, (eds), *Morphology. An International Handbook on Inflection and Word-formation*. Vol. 1. Berlin: Walter de Gruyter, 545–552.
- Morris, J., Frank, T., Grainger, J., and Holcomb, P. J. (2007). Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology* 44(4): 506–521.
- Moscoso del Prado Martín, F. (2003). *Paradigmatic Effects in Morphological Processing: Computational and Cross-linguistic Experimental Studies* (MPI Series in Psycholinguistics). Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.
- Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., and Baayen, R. H. (2004). Morphological family size in a morphologically rich language: the case of Finnish compared to Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition* 30: 1271–1278.
- Moscoso del Prado Martín, F., Deutsch, A., Frost, R., Schreuder, R., De Jong, N. H., and Baayen, R. H. (2005). Changing places: a cross-language perspective on frequency and family size in Hebrew and Dutch. *Journal of Memory and Language* 53: 496–512.
- Moscoso del Prado Martín, F., Ernestus, M., and Baayen, R. H. (2003). Do type and token effects reflect different mechanisms: Connectionist modelling of Dutch past-tense formation and final devoicing. *Brain and Language* 90: 287–298.
- Mourelatos, A. (1978). Events: processes and states. *Linguistics and Philosophy* 2(3): 415–434.
- Mous, M. (1993). *A Grammar of Iraqw* (Cushitic Language Studies 9). Hamburg: Helmut Buske.
- Mous, M. (2012). Cushitic. In Z. Frajzyngier and E. Shay (eds), *The Afroasiatic Languages*. Cambridge: Cambridge University Press, 342–422.
- Mufwene, S. S. (1984). The count/mass distinction and the English lexicon. In D. Testen, V. Mishra, and J. Drogo (eds), *Papers from the Parasession on Lexical Semantics*. Chicago, IL: Chicago Linguistic Society, 200–221.
- Mühleisen, S. (2010). *Heterogeneity in Word-Formation Patterns*. Amsterdam: Benjamins.
- Mulford, R. C. (1983). On the acquisition of derivational morphology in Icelandic: learning about *-ari*. *Íslenskt mál og almenn málfræði* 5: 105–125.
- Müller, P. O., Ohnheiser, I., Olsen, S., and Rainer, F. (eds) (forthcoming). *Word-Formation: An International Handbook of the Languages of Europe*. Berlin: Mouton De Gruyter.
- Munson, B. and Solomon, N. P. (2004). The effects of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research* 47: 1048–1058.
- Murphy, G. L. (2007). Parsimony and the psychological representation of polysemous words. In M. Rakova, G. Pethö, and C. Rákosi (eds), *The Cognitive Basis of Polysemy: New Sources of Evidence for Theories of Word Meaning*. Bern etc.: Lang, 47–70.
- Murre, J. M. J., Phaf, R. H., and Wolters, G. (1992). Calm: categorizing and learning module. *Neural Networks* 5: 55–82.
- Mushin, I. (2012). *A Grammar of (Western) Garrwa*, Pacific Linguistics 637. Berlin: De Gruyter Mouton.

- Mutaka, N. and Hyman, L. M. (1990). Syllable and morpheme integrity in Kinande reduplication. *Phonology* 7: 73–120.
- Muysken, P. (1988). Affix order and interpretation: Quechua. In M. Everaert, A. Evers, R. Huybregts, and M. Trommelen (eds), *Morphology and Modularity*. Dordrecht: Foris, 259–79.
- Nagano, A. (2007). Marchand's analysis of back-formation revisited: Back-formation as a type of conversion. *Acta Linguistica Hungarica* 54(1): 33–72.
- Nagano, A. (2010). Subject compounding and a functional change of the derivational suffix -ing in the history of English. In R. A. Cloutier, A. M. Hamilton-Brehm, and W. A. Kretzschmar, JR (eds), *Studies in the History of the English Language, V: Variation and Change in English Grammar and Lexicon: Contemporary Approaches* (Topics in English Linguistics: 68). Berlin: De Gruyter Mouton, 111–131.
- Nagano, A. (2011). The right-headedness of morphology and the status and development of category-determining prefixes in English. *English Language and Linguistics* 15: 61–83.
- Nagano, S. (1994). A note on the Tibetan kinship terms khu and zhang. *Linguistics of the Tibeto-Burman Area* 17(2): 103–115.
- Nagaraja, K. S. (1985). *Khasi: A Descriptive Analysis*. Pune: Deccan College Post-Graduate and Research Institute.
- Najlis, E. (1984). *Fonología de la protolengua mataguaya*. Buenos Aires: University of Buenos Aires.
- Narrog, H. and Heine, B. (2011). *The Oxford Handbook of Grammaticalization*. Oxford: Oxford University Press.
- Nash, D. (1986). *Topics in Warlpiri Grammar* (Outstanding Dissertations in Linguistics. Third Series) New York: Garland Publishing Inc.
- Nau, N. (1995). *Möglichkeiten und Mechanismen kontaktbewegten Sprachwandels unter besonderer Berücksichtigung des Finnischen* (Edition Linguistik, o8). Munich, Newcastle: LINCOM Europa.
- Nault, K. and Libben, G. (2004). Representation and processing of interfixed German verb-noun compounds. 4th International Conference on the Mental Lexicon, Windsor, Canada.
- Naumann, B. (1986). *Einführung in die Wortbildungslehre des Deutschen* (Germanistische Arbeitshefte, 4). Tübingen: Niemeyer.
- Naumann, B. and Vogel, P. M. (2000). Derivation. In G. Booij, C. Lehmann, and J. Mugdan (eds), *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*. Berlin: Walter de Gruyter, 929–942.
- Nedjalkov, I. (1994). Syntactic types of Evenki NPs. In F. Plank (ed.), *The Noun Phrase Sketch Book*. VII/22. EURO TYP working papers: 1–4.
- Nedjalkov, I. (1995). *Evenki*. London/Sydney/Dover/New Hampshire: Croom Helm.
- Nedjalkov, V. P. and Silnitsky, G. G. (1973). The typology of morphological and lexical causatives. In F. Kiefer (ed.), *Trends in Soviet Theoretical Linguistics*. Dordrecht: D. Reidel Publishing, 1–32.
- Neef, M. (1999). A declarative approach to conversion into verbs in German. In G. Booij and J. van Marle (eds), *Yearbook of Morphology 1998*. Dordrecht: Kluwer, 199–224.
- Neef, M. (forthcoming). Synthetic compounds in German. In P. Müller, I. Ohnheiser, S. Olsen, and F. Rainer (eds), *Word-Formation: An International Handbook of the Languages of Europe*. Berlin: Mouton de Gruyter.
- Nekushoeva, S. (2009). Pronominal adverbs in the Shughni language. Paper presented to the University of Kentucky Linguistics Program, March 2009.

- Nercesian, V. (2007). La nominalización de verbal y la estructura argumental heredada en wichí (mataco-mataguaya). Paper presented at the VI Encuentro de Lenguas Aborígenes y Extranjeras. Salta, Argentina.
- Nercesian, V. (2008). La negación en wichí (mataco-mataguaya). In C. Messineo, M. Malvestitti, and R. Bein (eds), *Estudios en lingüística y antropología. Homenaje a Ana Gerzenstein*. Buenos Aires: University of Buenos Aires, 167–189.
- Nercesian, V. (2009/10). Construcciones de verbos seriales en wichí (mataguaya). Características sintácticas y semánticas. *Amerindia* 33/34: 187–216.
- Nercesian, V. (2011a). Gramática del wichí, una lengua chaqueña. Interacción fonología-morfología-sintaxis en el léxico. Ph.D. dissertation. University of Buenos Aires.
- Nercesian, V. (2011b). Stress in Wichi (Mataguayan) and its interaction with the word formation processes. *Amerindia* 35: 75–102.
- Nercesian, V. (2012). Word-formation in polysynthetic languages: noun incorporation in Wichi (Mataguayan). Paper presented at Universals and Typology in Word-Formation II Conference, P. J. Šafárik University—SKASE, Košice, Slovakia, August 26–28, 2012.
- Nercesian, V. and Vidal, A. (forthcoming). Operaciones de aumento de valencia y clases verbales en wichí (mataguaya). In A. C. Bruno and F. Queixalós (eds), *Mudanças incrementais de valência nas línguas amaônicas*. Bogotá: University of Colombia.
- Nerlich, B. (2003). Polysemy: past and present. In B. Nerlich, Z. Todd, V. Herman, and D. D. Clarke (eds), *Polysemy: Flexible Patterns of Meaning in Mind and Language*. Berlin/New York: Mouton de Gruyter, 49–76.
- Nerlich, B. and Clarke, D. D. (2003). Polysemy and flexibility: introduction and overview. In B. Nerlich, Z. Todd, V. Herman, and D. D. Clarke (eds), *Polysemy. Flexible Patterns of Meaning in Mind and Language*. Berlin: Mouton de Gruyter, 3–30.
- Nevalainen, T. (1999). Early Modern English Lexis and Semantics. In R. Lass (ed.), *The Cambridge History of the English Language*, 3: 1476–1776. Cambridge: Cambridge University Press, 332–456. [First published 1993].
- Nevins, A. (2005). Overwriting does not optimize in nonconcanenative morphology. *Linguistic Inquiry* 36: 275–287.
- Nevins, A. and Vaux, B. (2003). Metalinguistic, shmetalinguistic: the phonology of shm-reduction. *Proceedings from the Regional Meetings. Chicago Linguistic Society* 39: 702–721.
- Nevis, J. and Joseph, B. (1993). Wackernagel affixes: Evidence from Balto-Slavic. *Yearbook of Morphology* 1992. Dordrecht: Kluwer Academic, 93–111.
- Newman, P. (2000). *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.
- Newman, P. (2007a). Hausa. In A. S. Kaye (ed.), *Morphologies of Asia and Africa*, vol. 1. Winona Lake, IN: Eisenbrauns, 677–711.
- Newman, P. (2007b). *A Hausa-English Dictionary*. Yale Language Series. Yale: Yale University Press.
- Newman, S. (1944). *Yokuts language of California*. New York: Viking Fund Publications.
- Newmeyer, F. J. (1986). *Linguistic Theory in America*. 2nd edn. New York: Academic Press.
- Ngunga, A. (2001). *Phonology and Morphology of the Ciyao Verb*. Stanford, CA: CSLI Publications.
- Nguyễn, Đình Hoà. (1966). *Vietnamese-English Dictionary*. Rutland, Vermont: Charles E. Tuttle Co.
- Nguyễn, Hữu Hoàn and Nguyễn Văn Lợi (1998). *Tiếng Katu* (The Katu language). Hanoi: Nhà Xuất Bản Khoa Học Xã Hội.

- Nguyễn, Hữu Hoàn (1995). *Katu language word formation*. Hanoi-Leiden Series: Data papers on minority languages of Vietnam. Hanoi: Nhà Xuất Bản Khoa Học Xã Hội.
- Nguyễn, Tài Cẩn (1996). *Ngữ pháp Tiếng Việt* (Vietnamese Grammar). Hanoi: Nhà Xuất Bản Đại Học Quốc Gia Hà Nội.
- Nguyễn, Văn Khang, Bùi Chỉ, and Hoàng Văn Hành. (2002). *Từ điển Mường-Việt* (A Muong-Vietnamese dictionary). Hanoi: Nhà Xuất Bản Văn Hoá Dân Tộc.
- Nguyễn, Văn Lợi (1993). *Tiếng Rục* (The Ruc language). Hanoi: Nhà Xuất Bản Khoa Học Xã Hội.
- Nguyễn, Văn Lợi, Nguyễn Hữu Hoàn, and Tạ Văn Thông (2008). *Tiếng Mảng* (The Mang Language). Nhà Xuất Bản Khoa Học Hà Nội.
- Nguyễn Văn Tài, Đoàn Văn Phúc, and Phan Xuân Thành (1984). *Sách học Tiếng Bru Vân Kiều* (A text to study the Bru language). Hanoi: Ủy Ban Nhân Dân, Tỉnh Bình Trị Thiên.
- Nguyễn Văn Tài, Đoàn Văn Phúc, and Phan Xuân Thành. (1986). *Sách học Tiếng Pakoh-Taôih* (Text for studying Pakoh-Taoh). Vietnam: Ủy Ban Nhân Dân, Tỉnh Bình Trị Thiên.
- Nichols, J. (1986). Head marking and dependent marking grammar. *Language* 62(1): 56–119.
- Nichols, J., Peterson, D., and Barnes, J. (2004). Transitivity and detransitivizing languages. *Linguistic Typology* 8: 149–212.
- Nida, E. A. (1948). The identification of morphemes. *Language* 24: 414–441.
- Niebuhr, O. and Kohler, K. J. (2011). Perception of phonetic detail in the identification of highly reduced words. *Journal of Phonetics* 39(3): 319–329.
- Niepokuj, M. K. (1997). *The Development of Verbal Reduplication in Indo-European*. Washington: Institute for the Study of Man.
- Nieuwenhuis, P. (1985). Diminutives. Ph.D. dissertation. University of Edinburgh.
- Nikolaeva, I. (2008). Between nouns and adjectives: a constructional view. *Lingua* 118: 969–996.
- Nikolaeva, I. and Tolskaya, M. (2001). *A Grammar of Udihe*. Berlin: Mouton de Gruyter.
- Nishida, T. (1994). A personal view of the Sino-Tibetan Language Family. *Current Issues in Sino-Tibetan Linguistics*. Osaka: 26th international conference on Sino-Tibetan Languages and Linguistics, 1–22.
- Noonan, M. P. (1997). Versatile nominalization. In J. Bybee, J. Haiman, and S. A. Thompson (eds), *Essays on Language Function and Language Type in Honor of Talmy Givón*. Amsterdam: Benjamins, 374–394.
- Norman, J. (1982). Four notes on Chinese-Altaic linguistic contacts. *Tsinghua Journal of Chinese Studies* 14(1–2): 243–7.
- Norris, D. G. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52: 189–234.
- Norris, D. and Kinoshita, S. (2008). Perception as evidence accumulation and bayesian inference: Insights from masked priming. *Journal of Experimental Psychology* 137(3): 434–455.
- Norris, D. and McQueen, J. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological Review* 115(2): 357–395.
- Nouguier Voisin, S. (2002). Relations entre fonctions syntaxiques et fonctions sémantiques en wolof. PhD thesis. University of Lyon.
- Novikova, K. et al. (1991). *Èvenskij jazyk* [The Even language]. Leningrad: Prosveščenie.
- Ntelitheos, D. (2007). Malagasy instrumental nominalizations. In D. S. Bigham, F. Hoyt, N. Seifert, A. Teodorescu, and J. White (eds), *The Proceedings of the Texas Linguistics Society IX Conference: The Morphosyntax of Under-Represented Languages*. Stanford, CA: CSLI Online Publications.
- Ntelitheos, D. (2012). *Deriving Nominals: A Syntactic Account of Malagasy Nominalizations*. Leiden: Brill Publications.

- Nuttall, P. A. (1840). *A Classical and Archæological Dictionary of the Manners, Customs, Laws, Institutions, Arts, etc. of the Celebrated Nations of Antiquity and of the Middle Ages*. London: Whittaker and Co.
- O'Grady, G. N. (1960). New concepts in Nyangumarda: some data on linguistic acculturation, *Anthropological Linguistics*, 2(1(January)): 1–6. Berlin: Mouton de Gruyter, 187–211.
- O'Grady, W. and de Guzman, V. P. (1996). Morphology: the analysis of word structure. In W. D. O'Grady, M. Dobrovolsky, and F. Katamba, (eds), *Contemporary Linguistics: An Introduction*. 3rd edn. London: Addison Wesley Longman Limited, 132–180.
- Ogura, M. (1995). The interchangeability of Old English Verbal Prefixes. *Anglo-Saxon England* 24: 67–93.
- Okada, A. (2010). The analyses of the English negative prefixes in the history of English. In O. Imahayashi, Y. Nakao, and M. Ogura (eds), *Aspects of the History of English Language and Literature* (Studies in English Medieval Language and Literature: 25). Frankfurt: Peter Lang, 343–351.
- Okell, J. and Allott, A. (2001). *Burmese/Myanmar Dictionary of Grammatical Forms*. Richmond: Curzon.
- Olsen, S. (1990). Konversion als kombinatorischer Wortbildungsprozeß. *Linguistische Berichte* 127: 185–216.
- Olsen, S. (1996). Partikelverben im deutsch-englischen Vergleich. In E. Lang and G. Zifonun (eds), *Deutsch-typologisch*. Berlin: Walter de Gruyter, 261–288.
- Olsen, S. (2000). Composition. In G. Booij, C. Lehman, and J. Mugdan (eds), *Morphology*. 1. Berlin: deGruyter, 897–916.
- Oltra-Massuet, I. (2010). On the morphology of complex adjectives. Ph.D. dissertation, Universitat Autònoma de Barcelona.
- Omar, A. H. (1976). The verb in Kentakbong. In P. N. Jenner, L. C. Thompson, and S. Starosta (eds), *Austroasiatic Studies Part II* (Oceanic Linguistics Special Publication No. 13). Honolulu: University of Hawaii Press, 951–960.
- ORCAWETA (Organización Capitanía Weenhayek y Tapiete) (2011). Resultados del censo Weenhayek. In *Plan Estratégico de desarrollo del pueblo weenhayek 2011–2015*. Bolivia: Tarija.
- O'Reilly, R. C. (1998). Six principles for biologically based computational models of cortical cognition. *Trends in Cognitive Science* 2: 455–462.
- O'Reilly, R. C. (2001). Generalization in interactive networks: the benefits of inhibitory competition and hebbian learning. *Neural Computation* 1199–1242.
- Orzechowska, A. (1999). Rzeczownik. In R. Grzegorzczkowska, R. Laskowski, and H. Wróbel (eds), *Gramatyka współczesnego języka polskiego: Morfologia*. Warszawa: PWN, 270–332.
- Osing, J. (1976). *Die Nominalbildung des Ägyptischen*. 2 vols (Sonderschriften des Deutschen Archäologischen Instituts, Abteilung Kairo). Mainz: von Zabern.
- Ott, D. (2011). Diminutive-formation in German. Spelling out the classifier analysis. *German Linguistics* 14: 1–46.
- Packard, J. L. (2000). *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge: Cambridge University Press.
- Palmer, B. (2009). *Kokota Grammar* (Oceanic Linguistics Special Publication no. 35). Honolulu: University of Hawai'i Press.
- Palmer, C. C. (2008). Borrowed derivational morphology in late Middle English: a study of the records of the London Grocers and Goldsmiths. In S. M. Fitzmaurice and D. Minkova (eds), *Studies in the History of the English Language, IV: Empirical and Analytical Advances*

- in the Study of English Language Change*. Topics in English Linguistics (TopicsEL): 61. Berlin: Mouton de Gruyter, 231–264.
- Palmer, C. C. (2009). Borrowings, derivational morphology, and perceived productivity in English, 1300–1600. Ph.D. dissertation. Ann Arbor, University of Michigan.
- Palsgrave, J. (1530). *L'claircissement de la langue française*. Suivi de la grammaire de Giles de Guez. Publiés pour la première fois en France, par F. Génin. Paris: Imprimerie nationale, 1852.
- Panagl, O. (1977). Aspekte der kindersprachlichen Wortbildung. In G. Drachman (ed.), *Salzburger Beiträge zur Linguistik: Akten der 3. Salzburger Jahresstagung für Linguistik* Salzburg: Neugebauer, 79–101.
- Panocová, R. (2009). On sound symbolism in evaluative morphology of East Slavonic Languages, *Sučasni doslidžennja z inozemnoji filolohiji*. Vypusk 7: zbirnyk naukovych prac. Užhorod: Užhorodskij nacional'nyj universytet, 324–328.
- Park, I. (2012). A Grammar of Hidasta. PhD thesis. Indiana University, Bloomington, IN.
- Parker, J. and Sims, A. D. (2012). Affix-ordering constraints and processing in Russian: A look at Complexity-based Ordering. Paper presented at the *American International Morphology Meeting* in Amherst, MA, September 21–23.
- Parkinson, S. (1988). Portuguese. In M. Harris and N. Vincent (eds), *The Romance Languages*. London: Croom Helm, 131–169.
- Parsons, T. (1990). *Events in the Semantics of English*. Cambridge, MA: MIT Press.
- Paster, M. (2006). Phonological conditions on affixation. Ph.D. dissertation. UC Berkeley.
- Paster, M. (2009). Explaining phonological conditions on affixation: evidence from suppletive allomorphy and affix ordering. *Word Structure* 2(1): 18–47.
- Pastizzo, M. J. and Feldman, L. B. (2009). Multiple dimensions of relatedness among words: conjoint effects of form and meaning in word recognition. *The Mental Lexicon* 4(1): 1.
- Patrie, J. (1982). *The Genetic Relationship of the Ainu Language*. Honolulu: University of Hawaii Press.
- Paul, H. (1880). *Prinzipien der Sprachgeschichte*. Halle: Niemeyer
- Paul, H. (1920). *Deutsche Grammatik*. Band V. Teil V: Wortbildungslehre. Tübingen: Niemeyer [reprint 1968].
- Paul, H. (1955). *Deutsche Grammatik*. Vol. 5: Wortbildungslehre. Halle: Niemeyer.
- Paul, H. (1968). *Deutsche Grammatik*. Bd. V. Teil V. Wortbildungslehre. Tübingen: Niemeyer.
- Paul, P. (1982). Homonyms, semantic divergence and valency. *Lingua* 58(3–4): 291–307.
- Pauly, É. (2010). *La polysémie*. Paris: L'Harmattan.
- Payne, T. E. (1997). *Describing Morphosyntax: A Field Guide for Linguists*. Cambridge: Cambridge University Press.
- Pearson, J. (1998). *Terms in Context*. Amsterdam: Benjamins.
- Pellegrino, C. C. and Marsico, E. (2011). Across-language perspective on speech information rate. *Language* 87(3): 539–558.
- Pelletier, F. J. (1975). Non-singular reference: some preliminaries. *Philosophia* 5(4): 451–465.
- Pengitov, N. T., Galkin, I. S., and Isanbaev, N. I. (eds) (1961). *Sovremennyy marijskij jazyk. Morfologija* [Contemporary Mari. Morphology]. Joškar-Ola: Marijskoe knižnoe izdatel'stvo.
- Pennanen, E. V. (1984). What happens in conversion? In H. Ringbom and M. Rissanen (eds), *Proceedings from the Second Nordic Conference for English Studies, Hanasaari/Hanaholmen, 19th–21st May 1983* (Meddelanden Från Stiftelsens för Åbo Akademi Forskningsinstitut 92). Åbo: Åbo Akademi, 79–93.

- Pentland, D. H. (1975). Diminutive consonant symbolism in Algonquian. *Papers of the Sixth Algonquian Conference* 1974, 237–252.
- Perea, M. and Carreiras, M. (2006). Do transposed-letter effects occur across lexeme boundaries? *Psychonomic Bulletin and Review* 13(3): 418–422.
- Perea, M. and Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language* 51(2): 231–246.
- Pereltsvaig, A. (2013). *Languages of the World: An Introduction*. Cambridge: Cambridge University Press.
- Pesetsky, D. (1995). *Zero Syntax: Experiencer and Cascades*. Cambridge, MA: MIT Press.
- Peters, H. (1991). Two Suffixes Reconsidered: Old English -ild and -incel. *English Studies: A Journal of English Language and Literature* 72(2): 106–122.
- Peters, H. (2006). The Old English Verbal Suffix -ettan. In A. J. Johnston, F. v. Mengden, and S. Thim (eds), *Language and Text: Current Perspectives on English and Germanic Historical Linguistics and Philology*. Anglistische Forschungen: 359. Heidelberg: Universitätsverlag Winter, 241–254.
- Peterson, D. (2007). *Applicative Constructions*. Oxford: Oxford University Press.
- Peterson, D. (2001). The Ingush clitic? a: the elusive type 5 clitic? *Language* 77: 144–155.
- Peterson, J. and Maas, U. (2009). Reduplication in Kharia: the masdar as a phonologically motivated category. *Morphology* 19: 207–237.
- Pétré, P. (2006). The Prefix be-/bi- as marker of verbs of deception in Late Old and Early Middle English. *Belgian Journal of English Language and Literature* 4: 109–127.
- Pétré, P. and Cuyckens, H. (2008). Bedusted, yet not beheaded: the role of be-s constructional properties in its conservation. In A. Bergs and G. Diewald (eds), *Constructions and Language Change* (Trends in Linguistics: Studies and Monographs: 194). Berlin: Mouton de Gruyter, 135–169.
- Pham, H. and Baayen, R. H. (2013). Semantic relations and compound transparency: a regression study in CARIN theory. *Psychologia* 46(4): 455–478.
- Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.
- Pinker, S. (1991). Rules of language. *Science* 253: 530–535.
- Pinker, S. (1997). Words and rules in the human brain. *Nature* 387: 547–548.
- Pinker, S. (1999). *Words and Rules*. New York: Harper Perennial.
- Pinnow, H. J. (1965). Personal pronouns in the Austroasiatic languages: a historical study. *Lingua* 14: 3–42.
- Pinnow, H. J. (1966). A comparative study of the verb in the Munda languages. In N. H. Zide (ed.), *Studies in Comparative Austroasiatic Linguistics*. The Hague: Mouton, 96–193.
- Piñón, C. (2001). A finer look at the causative-inchoative alternation. In *Proceedings of Semantics and Linguistic Theory* 11. Ithaca, NY: Cornell Linguistics Circle. References to prepublication manuscript on at: <<http://www.phil-fak.uni-duesseldorf.de/~pinon/papers/flcia.html>>.
- Pittman, C. (2006). Inuktitut Restructuring Affixes. *Proceedings of the Canadian Linguistics Association* <<http://westernlinguistics.ca/Publications/CLA2006/CLA-ACL2006.htm>>.
- Pittman, C. (2009). Complex verb formation revisited: restructuring in Inuktitut and Nuu-chah-nulth. In M.-A. Mahieu and N. Tersi (eds), *Variations on Polysynthesis: The Eskaleut Languages*. Amsterdam: Benjamins, 135–147.
- Plag, I. (1996). Selectional restrictions in English suffixation revisited: a reply to Fabb (1988). *Linguistics* 34: 769–798.

- Plag, I. (1999). *Morphological Productivity: Structural Constraints in English Derivation*. Berlin: Mouton de Gruyter.
- Plag, I. (2002). The role of selectional restrictions, phonotactics and parsing in constraining suffix order in English. *Yearbook of Morphology 2001*. Dordrecht: Kluwer Academic, 285–314.
- Plag, I. (2003). *Word-Formation in English*. Cambridge: Cambridge University Press.
- Plag, I. and Baayen, R. H. (2009). Suffix ordering and morphological processing. *Language* 85: 106–149.
- Plaisier, H. (2007). *A Grammar of Lepcha*. Leiden: Brill.
- Plank, F. (1981). *Morphologische (Ir-)Regularitäten*. Tübingen: Narr.
- Plank, F. (1994). Inflection and derivation. In R. E. Asher and J. M. Y. Simpson (eds), *Encyclopedia of Language and Linguistics*, Vol. 3. Oxford: Pergamon Press, 1671–1678.
- Plank, F. (1999). Split morphology: how agglutination and flexion mix. *Linguistic Typology* 3(3): 279–340.
- Plank, F. (2007). Extent and limits of linguistic diversity as the remit of typology—but through constraints on WHAT is diversity limited. *Linguistic Typology* 11: 43–68.
- Plank, F. (2010). Variable direction in zero-derivation and the unity of polysemous lexical items. *Word Structure* 3: 82–97.
- Plank, F., Filimonova, E., Mayorava, T., and Mayer, T. (n.d.). The universals archive. Retrieved from <<http://typo.uni-konstanz.de/archive/intro/>>.
- Platt, J. T. (1972). *An Outline Grammar of the Gugada Dialect: South Australia* (Australian Aboriginal Studies No. 48). Canberra: Australian Institute of Aboriginal Studies.
- Plaut, D. C. and Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes* 15(4/5): 445–485.
- Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005a). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica* 62: 146–159.
- Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005b). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America* 118: 2561–2569.
- Pollard, C. and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press and Stanford (CA): CSLI Publications.
- Pollatsek, A., Hyönä, J., and Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human, Perception and Performance* 26: 820–833.
- Pollock, J.-Y. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.
- Poppe, N. (1960). *Vergleichende Grammatik der altaischen Sprachen*. Teil I. Vergleichende Lautlehre. Wiesbaden: Otto Harrassowitz.
- Poser, W. (1989). The metrical foot in Diyari. *Phonology* 6: 117–148.
- Poser, W. (1990). Evidence for foot structure in Japanese. *Language* 66: 78–105.
- Poser, W. (2004). Noun classification in Carrier. Manuscript.
- Postal, P. M. (1969). Anaphoric islands. In Binnick et al. (eds), *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 209–239.
- Posti, L. and Itkonen, T. (eds) (1973). *FU-transkription yksinkertaistaminen. Az FU-átírás egyszerűsítése. Zur Vereinfachung der FU-Transkription. On simplifying of the FU transcription*. Castrenianumin toimitteita 7. Helsinki.

- Pounder, A. (2000). *Processes and Paradigms in Word-Formation Morphology*. Berlin: Mouton de Gruyter.
- Premisrat, S. (1987). *Khmu, a Minority Language of Thailand*. Pacific Linguistics A-76. Canberra, Australia: Australian National University.
- Priestley, J. (1762). *A course of lectures on the theory of the language, and universal grammar*. Warrington: W. Eyres.
- Prieto, V. M. (2005). Spanish evaluative morphology: pragmatic, sociolinguistic, and semantic Issues. Ph.D. dissertation. University of Florida.
- Prieto, V. M. (forthcoming). The semantics of evaluative morphology. In N. Grandi and L. Körtvélyessy (eds), *Edinburgh Handbook of Evaluative Morphology*. Edinburgh: Edinburgh University Press.
- Prince, A. and Smolensky, P. (1993). Optimality theory: constraint interaction in generative grammar. MS. New Brunswick: Rutgers University, Cognitive Science Center.
- Prince, E. F. (1998). The borrowing of meaning as a cause of internal syntactic change. In M. S. Schmid, J. R. Austin, and D. Stein (eds), *Historical Linguistics 1997: Selected Papers from the 13th International Conference on Historical Linguistics*, Düsseldorf, 10–17.
- Pulleyblank, D. (1986). *Tone in Lexical Phonology*. Dordrecht: Reidel.
- Pulleyblank, E. G. (1973). Some new hypotheses concerning word families in Chinese. *Journal of Chinese Linguistics* 1(1): 111–125.
- Pullum, G. and Zwicky, A. (1999). Gerund participles and head-complement inflection conditions. In P. Collins and D. Lee (eds), *The Clause in English: In Honour of Rodney Huddleston*, Amsterdam: Benjamins, 251–271.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Puura, U. (2010). Frequentative and momentative verbal derivation in Veps. *Linguistica Uralica* 46(4): 261–280.
- Pye, C. (1986). Quiché Mayan speech to children. *Journal of Child Language* 13: 85–100.
- Pylkkänen, L. (2002). Introducing arguments. Ph.D. dissertation. MIT, Cambridge, MA.
- Pylkkänen, L., Feintuch, S., Hopkins, E., and Marantz, A. (2004). Neural correlates of the effects of morphological family frequency and family size: an MEG study. *Cognition* 91: B35–B45.
- Pylkkänen, L., R. Llinás, and G. L. Murphy (2006). The representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience* 18: 97–109.
- Quirk, R. and Wrenn, C. L. (1958). *An Old English grammar*. 2nd edn. London: Methuen.
- Raa, E. ten, Green, W. and Tod, S. (1974). Research dictionary of the Western Desert language of Australia, Part I: vernacular—English. Manuscript. Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies.
- Rabin, J. and Deacon, H. (2008). The representation of morphologically complex words in the developing lexicon. *Journal of Child Language* 35: 453–465.
- Rackowski, A. (1999). Morphological optionality in Tagalog aspectual reduplication. *Papers on Morphology and Syntax, Cycle Two, MITWPL* 34. Cambridge: MITWPL, 107–136.
- Radhakrishnan, R. (1976). A note on the morphology of the causative in Nancowry. In Philip N. Jenner et al. (ed.), *Austroasiatic Studies*. Honolulu: The University Press of Hawaii, 1035–1040.
- Radhakrishnan, R. (1981). The Nancowry word: phonology, affixal morphology and roots of a Nicobarese language. In *Current Inquiry into Language and Linguistics* 37. Edmonton, Alberta, Canada: Linguistic Research, Inc.
- Rainer, F. (1988). Towards a theory of blocking: the case of Italian and German quality nouns. In G. Booij and J. van Marle (eds), *Yearbook of Morphology 1988*. Dordrecht: Foris, 155–185.

- Rainer, F. (1996). La polysémie des noms abstraits: historique et état de la question. In N. Flaux, M. Glatigny and D. Samain (eds), *Les noms abstraits. Histoire et théories*. Villeneuve d'Ascq: Presses Universitaires du Septentrion, 117–126.
- Rainer, F. (1999). La derivación adjetival. In I. Bosque and V. Demonte (eds), *Gramática descriptiva de la lengua española*. Madrid: Espasa, 4595–4643.
- Rainer, F. (2003). Semantic fragmentation in word-formation: the case of Spanish -azo. In R. Singh and S. Starosta (eds), *Explorations in Seamless Morphology*. New Dehli: Sage Publications, 197–211.
- Rainer, F. (2005a). Semantic change in word formation, *Linguistics* 43(2): 415–441.
- Rainer, F. (2005b). Typology, diachrony, and universals of semantic change in word-formation: Romanist's look at the polysemy of agent nouns. In G. Booij, E. Guevara, A. Ralli, S. Sgroi, and S. Scalise (eds), *Morphology and Linguistic Typology*. On-line Proceedings of the Fourth Mediterranean Morphology Meeting (MMM4), 21–34.
- Rainer, F. (2005c). Constraints on productivity. In P. Štekauer and R. Lieber (eds), *Handbook of Word-Formation*. Dordrecht: Springer, 335–352.
- Rainer, F. (2010). Sobre polisemia en la formación de palabras, *Hesperia* 13(2): 7–52.
- Rainer, F. (2011). The agent-instrument-place 'polysemy' of the suffix -TOR in Romance, *Language Typology and Universals* 64(1): 8–32.
- Rainer, F. and Varela, S. (1992). Compounding in Spanish. *Rivista di linguistica* 4(1): 117–142.
- Ramat, A. G. and Hopper, P. J. (1998). Introduction. In A. Giacalone Ramat and P. J. Hopper (eds), *The Limits of Grammaticalization*. Amsterdam: Benjamins, 1–10.
- Ramat, P. (1998a). The Germanic languages. In A. Giacalone Ramat and P. Ramat (eds), *The Indo-European Languages*. London: Routledge, 380–414.
- Ramat, P. (1998b). Typological comparison and linguistic areas: some introductory remarks. *Language Sciences* 20(3): 227–40.
- Ramchand, G. C. (2008). *Verb Meaning and the Lexicon: A First-phase Syntax*. Cambridge: Cambridge University Press.
- Ramisch, H. (1989). *The Variation of English in Guernsey/Channel Islands*. (Bamberger Beiträge zur Englischen Sprachwissenschaft, 24). Frankfurt: Peter Lang.
- Ramscar, M. and Baayen, R. (2013). Production, comprehension and synthesis: A communicative perspective on language. *Frontiers in Psychology* 4: 233.
- Ramscar, M. and Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends In Cognitive Science* 11(7): 274–279.
- Ramscar, M. and Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: a new approach to the logical problem of language acquisition. *Cognitive Science* 31(6): 927–960.
- Ramscar, M., Dye, M., Gustafson, J., and Klein, J. (2013). Dual routes to cognitive flexibility: Learning and response conflict resolution in the dimensional change card sort task. *Child Development* 84(4): 1308–1323.
- Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., and Baayen, R. (2014). Nonlinear dynamics of lifelong learning: the myth of cognitive decline. *Topics in Cognitive Science* 6: 5–42.
- Ramstedt, G. J. (1952). *Einführung in die altaische Sprachwissenschaft II. Formenlehre*, edited and published by Pentti Aalto. Helsinki: Suomalais-Ugrilainen Seura.
- Ramstedt, G. J. (1957). *Einführung in die altaische Sprachwissenschaft I. Lautlehre*, edited and published by Pentti Aalto. Helsinki: Suomalais-Ugrilainen Seura.
- Ramstedt, G. J. (1966). *Einführung in die altaische Sprachwissenschaft III. Register*, edited and published by Pentti Aalto. Helsinki: Suomalais-Ugrilainen Seura.

- Rappaport Hovav, M., and Levin, B. (1992). *-er* Nominals: implications for a theory of argument structure. In T. Stowell and E. Wehrli (eds), *Syntax and Semantics 26: Syntax and the Lexicon*. New York: Academic Press, 127–153.
- Rappaport Hovav, M. and Levin, B. (1996). Two types of dreived accomplishments. In M. Butt and T. King (eds), *Proceedings of the First LFG Conference*. Grenoble: RANK Xerox, 375–388.
- Rappaport Hovav, M. and Levin, B. (1998). Building verb meaning. In M. Butt and W. Geuder (eds), *The Projection of Arguments: Lexical and Compositional Factors*. Stanford, CA: CSLI Publications, 97–134.
- Rappaport Hovav, M. and Levin, B. (2010). Reflections on Manner/Result Complementarity. In M. Rappaport Hovav, E. Doron, and I. Sichel (eds), *Syntax, Lexical Semantics, and Event Structure*. Oxford: Oxford University Press, 21–38.
- Rastle, K. and Davis, M. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes* 23(7–8): 942–971.
- Rastle, K., Davis, M. H., and New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review* 11: 1090–1098.
- Ratcliffe, R. R. (2005). Bukhara Arabic: a metatypized dialect of Arabic in Cental Asia. In E. A. Csató, B. Isaksson, and C. Jahani (eds), *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*. London: Routledge Curzon, 141–59.
- Rau, D.-H. V. (1992). *A Grammar of Atayal*. Ph.D. dissertation. Cornell University.
- Rau, D. V. and Dong, M.-N. (2005). *Concentric: Studies in Linguistics* 31(2): 57–87.
- Rayfield, J. R. (1970). *The Language of a Bilingual Community* (Janua Linguarum, Series Practica, 77). The Hague: Mouton.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124(3): 372–422.
- Refsing, K. (1986). *The Ainu Language: The Morphology and Syntax of the Shizunai Dialect*. Århus: Aarhus University Press.
- Regamey, C. (1946/1947). Considérations sur le système morphologique du tibétain littéraire. *Cahiers Ferdinand de Saussure* 6: 26–46.
- Regier, T. (1994). A preliminary study of the semantics of reduplication. TR-94-019. International Computer Science Institute, Berkeley, CA.
- Reh, M. (1996). *Anywa Language: Description and Internal Reconstructions*. Cologne: Rüdiger Köppe.
- Reichard, G. (1959). A comparison of five Salish languages. *International Journal of American Linguistics* 25: 239–253.
- Reid, L. A. (1999). New linguistic evidence for the Austric hypothesis. In E. Zeitoun and P. J.-K. Li (eds), *Selected Papers from the Eighth International Conference on Austronesian Linguistics*. Taipei: Academia Sinica, 5–30.
- Reid, L. A. (2009). On the diachronic development of C₁V₁-reduplication in some Austronesian languages. *Morphology* 19: 239–261.
- Reinisch, L. (1895). *Wörterbuch der Bedäuye-sprache*. Wien: Alfred Hölder.
- Reintges, C. (1994). Egyptian root-and-pattern morphology. *Lingua Aegyptia* 4: 213–244.
- Revel-Macdonald, N. (1979). *Le Palawan (Philippines): phonologie, catégories, morphologie* (Langues et civilisations de l'Asie du sud-est et du monde insulindien, No. 4). Paris: SELAF.
- Rice, K. (1989). *A Grammar of Slave*. Berlin: Mouton de Gruyter.

- Rice, K. (2000). *Morpheme Order and Semantic Scope: Word Formation in the Athapaskan Verb*. Cambridge: Cambridge University Press.
- Richards, E. and Hudson, J. (1990). *Walmajarri-English dictionary with English finder list*. Darwin: Summer Institute of Linguistics.
- Riddle, E. (1985). A Historical Perspective on the Productivity of the Suffixes -ness and -ity. In J. Fisiak (ed.), *Historical Semantics/ Historical Word-Formation*. Berlin: Mouton, 435–461.
- Riese, T. (2001). *Historische Nominalderivation des Wogulischen*. Wiesbaden: Harrassowitz.
- Rischel, J. (1974). *Topics in West Greenlandic Phonology: Regularities Underlying the Phonetic Appearance of Wordforms in a Polysynthetic Language*. Copenhagen: Akademisk Forlag.
- Rischel, J. (1995). *Minor Mlabri: A Hunter-gatherer Language of Northern Indochina*. Denmark: Museum of Tusculanum Press, University of Copenhagen.
- Robbeets, M. (2004). Swadesh 100 on Japanese, Korean and Altaic. *Tokyo University Linguistic Papers*, TULIP 23: 99–118.
- Robbeets, M. (2005). *Is Japanese related to Korean, Tungusic, Mongolic and Turkic?* Wiesbaden: Otto Harrassowitz.
- Robbeets, M. (2007). How the actional suffix chain connects Japanese to Altaic. *Turkic Languages* 11(1): 3–58.
- Roberts, I. (2003). Head movement. In M. Baltin and C. Collins (eds), *Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, 113–147.
- Roberts, J. (1987). *Amele*. London: Croom Helm.
- Roberts, J. (1991). Reduplication in Amele. *Papers in Papuan Linguistics* 1: 115–146.
- Robin, F. and Vittrant, A. (2007). Reduplication dans les langues tibeto-birmanes. L'exemple du birman et du tibétain. *Faits de langues* 29: 77–98.
- Robins, R. H. (1958). *The Yurok Language: Grammar, Texts, Lexicon*. Berkeley: University of California Press.
- Robins, R. H. (1979). *A Short History of Linguistics*. 2nd edn. London: Longman.
- Robinson, J. O. (1976). His and hers morphology: the strange case of Tarok possessives. In L. Hyman, L. Jacobson and R. Schuh (eds), *Studies in African Linguistics* (Supplement 6). Los Angeles: Department of Linguistics, UCLA, 201–209.
- Rodrigues, A. S. (2009). Portuguese converted de-verbal nouns: constraints on their bases. *Word Structure* 2: 69–107.
- Roelofs, A. (1997). Morpheme frequency in speech production: testing WEAVER. In G. E. Booij and J. Van Marle (eds), *Yearbook of Morphology 1996*. Kluwer: Dordrecht, 135–154.
- Roelofs, A. and Baayen, R. H. (2002). Morphology by itself in planning the production of spoken words. *Psychonomic Bulletin and Review* 9: 132–138.
- Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (eds), *Cognition and Categorization*. Hillside, NJ: Lawrence Erlbaum, 27–48.
- Rose, S. (1994). Palatalization, underspecification, and plane conflation in Chaha. *Proceedings of the West Coast Conference on Formal Linguistics* 12: 101–116.
- Rose, S. (1997). Theoretical issues in comparative ethio-Semitic phonology and morphology. Ph.D. thesis. McGill University, Montreal, Canada.
- Ross, J. R. (1972). The category squish: Endstation Hauptwort. In P. M. Peranteau, J. N. Levi, G. C. Phares, et al. (eds), *Proceedings of the Eighth Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, University of Chicago, Chicago, Illinois, 316–338.
- Ross, M. D. (1996). Contact-induced change and the comparative method: cases from Papua New Guinea. In M. Durie and M. D. Ross (eds), *The Comparative Method Reviewed:*

- Regularity and Irregularity in Language Change*. New York: Oxford University Press, 180–217.
- Ross, M. D. (2007). Calquing and metatypy. *Journal of Language Contact* 1: 116–143.
- Rothmayr, A. (2009). *The Structure of Stative Verbs*. Amsterdam: Benjamins.
- Rothstein, S. (2004). *Structuring Events. A Study in the Semantics of Lexical Aspect*. Oxford: Blackwell.
- Roy, C. and Labelle, M. (2007). Connaissance de la morphologie dérivationnelle chez les francophones et non-francophones de 6 à 8 ans. *Canadian Journal of Applied Linguistics, Revue Canadienne de Linguistique Appliquée* 10: 263–291.
- Roy, I. (2010). De-adjectival nominalizations and the structure of the adjective. In A. Alexiadou and M. Rathert (eds), *The Syntax of Nominalizations across Languages and Frameworks*. Berlin: Mouton de Gruyter, 129–158.
- Rubin, G. and Turano, K. (1992). Reading without saccadic eye movements. *Vision Research* 32(5): 895–902.
- Rubino, C. R. G. (2000). *Ilocano Dictionary and Grammar*. PALI Language Texts. Honolulu: University of Hawaii Press.
- Rubino, C. (2005). Reduplication: form, function and distribution. In B. Hurch and V. Mattes (eds), *Studies on Reduplication*. Berlin: Mouton de Gruyter, 11–30.
- Rubino, C. (2011). Reduplication. In M. S. Dryer and M. Haspelmath (eds), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 27. Available online at <<http://wals.info/chapter/27>> accessed on August 1, 2012.
- Rueckl, J. G. and Rimzhim, A. (2011). On the interaction of letter transpositions and morphemic boundaries. *Language and Cognitive Processes* 26(4–6): 482–508.
- Ruhl, C. (1989). *On Monosemy. A Study in Linguistic Semantics*. Albany, NY: State University of New York Press.
- Ruog, I. (1943). *Lappische Verbalableitung dargestellt auf Grundlage des Pitelappischen*. Uppsala/Leipzig: Lundequist/Harrassowitz.
- Ryan, K. M. (2010). Variable affix order: Grammar and learning. *Language* 86(4): 758–791.
- Ryder, M. E. (1994). *Ordered Chaos: The Interpretation of English Noun-Noun Compounds*. Berkeley: University of California Press.
- Ryder, M. E. (1999). Bankers and blue-chippers: an account of -er formations in Present-day English. *English Language and Linguistics* 3(2): 269–297.
- Saciuk, B. (1969). The stratal division of the lexicon. *Papers in Linguistics* 1: 464–532.
- Sadock, J. (1980). Noun incorporation in Greenlandic: A case of syntactic word formation. *Language* 56: 300–319.
- Sadock, J. (1986). Some notes on noun incorporation. *Language* 62(1): 19–31.
- Sadock, J. (2003). *A Grammar of Kalaallisut (West Greenlandic Inuttut)*. Languages of the World/Materials 162. Munich: LINCOM EUROPA.
- Saeed, J. I. (1997). *Semantics*. Oxford: Blackwell.
- Saeed, J. I. (2007). Somali morphology. In A. S. Kaye (ed.), *Morphologies of Asia and Africa*, vol. 1. Winona Lake, IN: Eisenbrauns, 547–586.
- Sagart, L. (1999). *The Roots of Old Chinese*. Amsterdam: Benjamins.
- Sagart, L. (2006). On intransitive nasal prefixation in Sino-Tibetan languages. *Cahiers de linguistique—Asie orientale* 35(1): 57–70.
- Säily, T. (2011). Variation in morphological productivity in the BNC: Sociolinguistic and methodological considerations. *Corpus Linguistics and Linguistic Theory* 7(1): 119–141.
- Sakel, J. (2004). *A Grammar of Mosetén*. München: Mouton de Gruyter.

- Samek-Lodovici, V. (1992). A unified analysis of cross-linguistic morphological gemination. *Proceedings of ConSole* 1: 265–283.
- Sammallahti, P. (1998). *The Saami Languages: An Introduction*. Kárášjohka: Davvi Girji.
- Sanchez, T. S. (2005). Constraints on structural borrowing in a multilingual contact situation. Ph.D. dissertation. University of Pennsylvania.
- Sander, P. and Sanders, L. (2006). Rogue males: sex differences in psychology students. *Electronic Journal of Research in Educational Psychology* 8(4): 1.
- Sanders, G. (1988). Zero derivation and the Overt Analogue Criterion. In M. Hammond and M. Noonan (eds), *Theoretical Morphology. Approaches in Modern Linguistics*. San Diego, CA: Academic Press, 155–175.
- Sapir, D. J. (1965). *A Grammar of Diola-Fogny. A Language Spoken in the Basse-Casamance Region of Senegal*. Cambridge: Cambridge University Press.
- Sapir, E. (1911). The problem of noun incorporation in American languages. *American Anthropologist* 13: 250–282.
- Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Co.
- Sapir, E. (1930). The Southern Paiute language. *Proceedings of the American Academy of Arts and Sciences* 65: 1–296.
- Sapir, E. (1944). Grading: A study in semantics, *Philosophy of Science* 11: 93–116.
- Sauer, G. (1967). *Die Nominalbildung im Ostjakischen*. Berlin: Akademie-Verlag.
- Saussure, F. de (1959). *Course in General Linguistics* (edited by C. Bally and A. Sechehaye, translated by W. Baskin). New York: McGraw Hill.
- Savà, G. (2005). *A Grammar of Tšamakko*. Kuschischen Sprachstudien 22. Cologne: Rüdiger Köppe.
- Savickienė, I. and Dressler, W. U. (2007). *The Acquisition of Diminutives*. Amsterdam: Benjamins.
- Saxton, D. (1963). Papago phonemes. *International Journal of American Linguistics* 29: 29–35.
- Scalise, S. (1984). *Generative Morphology*. Dordrecht: Foris.
- Scalise, S. (1988). Inflection and derivation. *Linguistics* 26: 561–581.
- Scalise, S. (1992). Compounding in Italian. *Rivista di linguistica* 4(1): 175–198.
- Scarborough, R. A. (2004). Coarticulation and the structure of the lexicon. Ph.D. dissertation. UCLA.
- Scarpa, G. F. (2007). Hacia una etnotaxonomía vegetal Chorote I: Fitonimia, sistema nomenclatural y comparación dialectal. *Suplemento Antropológico* 42(1): 81–119.
- Scarpa, G. F. (2010). Hacia una etnotaxonomía vegetal chorote II. Clasificación de las plantas entre las parcialidades *iyojwa'ja* y *iyowujwa* del Chaco argentino. In C. Messineo, G. Scarpa, and F. Tola (eds), *Léxico y categorización etnobiológica en grupos indígenas del Gran Chaco*. Santa Rosa: University of La Pampa, 157–198.
- Schachter, P. and Fe Otones (1972). *Tagalog Reference Grammar*. Berkeley: University of California Press.
- Schadeberg, T. C. (2003). Derivation. In D. Nurse and G. Philippson (eds), *The Bantu Languages*. London: Routledge, 71–89.
- Schaeffer, R. P. and F. O. Ebgokhare (2002). On the status of DO/SAY verbs with Emai ideophones. *Anthropological Linguistics* 44(3): 278–296.
- Schaerlakens, A. M. (1980). *De taalontwikkeling van het kind*. 2nd edn. Groningen: Wolters-Noordhoff.
- Schäfer, F. (2008). *The Syntax of (Anti-)Causatives: External Arguments in Change-of-state Contexts*. Amsterdam: Benjamins.

- Schäfer, F. (2009). The causative alternation. *Language and Linguistics Compass* 3: 641–681.
- Schäfer, F. (2011). Naturally atomic er-nominalizations. *Recherches linguistiques de Vincennes* 40: 149–174.
- Schebeck, B. (1998). Yura Combined Dictionary. Unpublished, electronic dictionary.
- Schebesta, P. P. (1926). Grammatical sketch of the Jahai dialect, spoken by a Negrito tribe of the Ulu Perak and Ulu Kelantan, Malay Peninsula. *The Bulletin of the School of the Oriental Society XX BSOAS* 4: 803–826.
- Schebesta, P. P. (1931). Grammatical sketch of the Ple-Temer language. *Journal of the Royal Asiatic Society of Great Britain and Ireland* 1931: 641–652.
- Schilling, G. (2007). *De jacht of Planeet X: Sterrenkundigen ontdekken de buitendelen van het zonnestelsel*. 's-Gravenland: Fontaine.
- Schmid, H.-J. (2011). *English Morphology and Word-Formation*. 2nd rev and trans edn. Berlin: Schmidt.
- Schmid, M. S., Austin, J. R., and Stein, D. (eds) (1998). *Historical Linguistics 1997*. Selected papers from the 13th International Conference on Historical Linguistics, Düsseldorf, August 10–17, 1997. (Amsterdam Studies in the Theory and History of Linguistic Science, 164). Amsterdam: Benjamins.
- Schmidt, W. (1906). Die Mon-Khmer-Völker: Ein Bindeglied zwischen Völkern Zentralasiens und Austronesians. *Archiv für Anthropologie* 33: 5.
- Schmidt, W. (1907). Die Mon-Khmer-Völker. Ein Bindeglied zwischen Völkern Zentralasiens und Austronesians. Braunschweig: Vieweg. [Translated as *Les peuples monkhmer. Trait d'union entre les peuples de l'Asie centrale et de l'Austronesie* (Bulletin de l'Ecole Française d'Extrême-Orient, 7–8; Hanoi, 1907).]
- Schneider, E. W. (1988). *Variabilität, Polysemie und Unschärfe der Wortbedeutung*. Vol. 1. *Theoretische und methodische Grundlagen*. Tübingen: Niemeyer.
- Schneider, K. (2003). *Diminutives in English*. Tübingen: Niemeyer.
- Schreuder, R. and Baayen, R. H. (1994). Prefix-stripping re-revisited. *Journal of Memory and Language* 33: 357–375.
- Schreuder, R. and Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (ed.), *Morphological Aspects of Language Processing*. Hillsdale, NJ: Lawrence Erlbaum, 31–154.
- Schreuder, R. and Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language* 37: 118–139.
- Schröder, A. (2008). Investigating the morphological productivity of verbal prefixation in the history of English. *Arbeiten aus Anglistik und Amerikanistik* 33(1): 47–69.
- Schuessler, A. (1985). The function of Qusheng in early Zhou Chinese. In G. Thurgood et al. (eds), *Linguistics of the Sino-Tibetan area: The State of the Art* (Pacific Linguistics series C-No.87). Canberra: The Australian National University, 344–362.
- Schuessler, A. (1998). Another note on Old Tibetan rje-blas. *Linguistics of the Tibeto-Burman Area* 21(2): 3–4.
- Schuh, R. (1998). *A Grammar of Miya*. Berkeley, Los Angeles, and London: University of California Press.
- Schuh, R. (2007). Bade morphology. In A.S. Kaye (ed.), *Morphologies of Asia and Africa*, vol. 1. Winona Lake, IN: Eisenbrauns, 587–639.
- Schultink, H. (1961). Produktiviteit als morfologisch fenomeen. *Forum der Letteren* 2: 110–125.

- Schultze-Berndt, E. (2000). Simple and complex verbs in Jaminjung: a study of event categorisation in an Australian languages. Ph.D. dissertation. Katholieke Universiteit Nijmegen.
- Schuppler, B., van Dommelen, W. A., Koreman, J., and Ernestus, M. (2012). How linguistic and probabilistic properties of a word affect the realization of its final/t: studies at the phonemic and sub-phonemic level. *Journal of Phonetics* 40: 595–607.
- Schürmann, C. W. (1844). *A Vocabulary of the Parnkalla Language. Spoken by the Natives Inhabiting the Western Shores of Spencer's Gulf. To Which Is Prefixed a Collection of Grammatical Rules, Hitherto Ascertained*. Adelaide: George Dehne.
- Scorza, D. P. (1985). A sketch of AU morphology and syntax. *Pacific Linguistics A* 63.
- Segalowitz, S. and Zheng, X. (2009). An ERP study of category priming: evidence of early lexical semantic access. *Biological Psychology* 80(1): 122.
- Seidenberg, M. (1987). Sublexical structures in visual word recognition: access units or orthographic redundancy. In M. Coltheart (ed.), *Attention and Performance XII*, Hillsdale NJ: Lawrence Erlbaum Associates, 245–264.
- Seidenberg, M. S. and Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences* 4(9): 353–361.
- Seidler, S. (1988). Untersuchung zum Erwerb von Wortbildungsregeln Deverbativa Nomina (Agens und Instrument) im Französischen. Unpublished MA thesis, Universität Hamburg.
- Selkirk, E. O. (1982). *The Syntax of Words*. Cambridge, MA: MIT Press.
- Sells, P. (1995). Korean and Japanese morphology from a lexical perspective. *Linguistic Inquiry* 277–325.
- Serbat, G. (1975). *Les dérivés nominaux latins à suffixe médiatif*. Paris: Les Belles Lettres.
- Shafer, R. (1938). Prefixed m- in Tibetan. *Sino-Tibetica* 3: 11–28.
- Shapiro, M. C. (2003). Hindi. In G. Cardona and D. Jain (eds), *The Indo-Aryan Languages*. London: Routledge, 250–285.
- Sharp, J. (2004). *Nyangumarta: A Language of the Pilbara region of Western Australia* (Pacific Linguistics; 556). Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies.
- Sharpe, M. C. (1995). *Dictionary of Western Bundjalung*. Armidale NSW: The Author.
- Sharpe, M. C. (1996). *An Introduction to the Yugambeh-Bundjalung language and its Dialects*. 3rd edn. Armidale: The Author.
- Shay, E. (1999). A grammar of East Dangla: the simple sentence. Ph.D. dissertation. University of Colorado. Ann Arbor: UMI.
- Shay, E. (in progress). *A Grammatical Sketch of Pévé*.
- Shay, E. (submitted). *A Grammar of Giziga*.
- Shibatani, M. (1990). *The Languages of Japan*. Cambridge: Cambridge University Press.
- Short, D. (2007). *Family Tree of Indo-European Languages*. <<http://www.danshort.com/ie/>> accessed December 4, 2011.
- Shorto, H. L. (1963). The structural patterns of Northern Mon-Khmer languages. *Linguistic Comparison of Southeast Asia and the Pacific*. University of London: School of Oriental and African Studies, 47–61.
- Shorto, H. L., Sidwell, P., Cooper, D. and Bauer, C. (eds) (2006). *A Mon-Khmer Comparative Dictionary*. Canberra: Australian National University, Pacific Linguistics.
- Sibanda, G. (2004). Verbal phonology and morphology of Ndebele. Ph.D. dissertation. University of California, Berkeley.
- Sidwell, P. (2008). Issues in the morphological reconstruction of Proto-Mon-Khmer. In C. Bowern, B. Evans, and L. Miceli (eds), *Morphology and Language History: In Honour of Harold Koch* 251–265.

- Siegel, D. (1974). Topics in English morphology. Ph.D. dissertation. MIT.
- Siegel, D. (1979). *Topics in English Morphology*. New York: Garland.
- Siewierska, A. (1984). *The Passive: A Comparative Linguistic Analysis*. London: Croom Helm.
- Siirainen, M. (2005). An overview of mental verbs in Finnish. In M. M. Jocelyne Fernandez-Vest (ed.), *Les langues ouraliennes aujourd'hui. Approche linguistique et cognitive* (The Uralic languages today: a linguistic and cognitive approach) (Bibliothèque de l'Ecole des hautes études. Sciences historiques et philologiques, 340). Paris: Champion 316–326.
- Silver, S. and Miller, W. R. (1997). *American Indian Languages—Cultural and Social Contexts*. Tucson: The University of Arizona Press.
- Simon, W. (1941). Certain Tibetan suffixes and their combinations. *Harvard Journal of Asiatic Studies* 5: 388–389.
- Simpson, J. (1985). How Warumungu people express new concepts. *Language in Central Australia* 4: 12–25.
- Simpson, J. (1991). *Warlpiri Morphosyntax: A Lexicalist Approach* (Studies in Natural Language and Linguistic Theory). Dordrecht: Kluwer.
- Simpson, J. (2001). Preferred word order, argument structure and the grammaticalisation of associated path. In M. Butt and T. H. King (eds), *Time over Matter: Diachronic Perspectives on Morphosyntax*. Stanford, CA: CSLI Publications, 173–208.
- Simpson, J. (2008a). Language landscapes of children in remote Australia. In R. Amery and J. Nash (eds), *Warra wiltaniappendi: Strengthening Languages: Proceedings of the Inaugural Indigenous Languages Conference (ILC)*, September 24–27, 2007. Adelaide: University of Adelaide, 51–60.
- Simpson, J. (2008b). Warumungu vocabulary. Unpublished electronic file. Canberra.
- Simpson, J. (2009). Warlpiri compounds. In R. Lieber and P. Štekauer (eds), *Oxford Handbook of Compounding*. Oxford: Oxford University Press, 609–622.
- Simpson, J. and Withgott, M. (1986). Pronominal clitic clusters and templates. In H. Borer (ed.), *Syntax and Semantics 19: The Syntax of Pronominal Clitics*. New York: Academic, 149–174.
- Simpson, J. A. (1989). *The Oxford English Dictionary*. Oxford: Clarendon Press.
- Singh, R. (1982). On some 'redundant compounds' in Modern Hindi. *Lingua* 56: 345–351.
- Singh, R. (1998). *Linguistic Theory, Language Contact, and Modern Hindustani: The Three Sides of a Linguistic Story*. New Delhi: Munshiram Manoharlal Publishers Pvt Ltd.
- Singh, R. (2005). Reduplication in Modern Hindi and the theory of reduplication. In B. Hurch (ed.), *Studies on Reduplication*. Berlin: Mouton de Gruyter, 263–282.
- Singh, R. and Agnihotri, R. K. (1997). *Hindi Morphology: A Word-based Description*. Delhi: Motilal Banarsidass.
- SKJa (1955). *Sovremennyy komi jazyk I. Fonetika, leksika, morfologija*. [Contemporary Komi language I: Phonetics, lexicon, morphology.] Syktyvkar: SKJa.
- Slater, K. W. (2003). *A Grammar of Mangghuer*. London: Routledge Curzon.
- Slobin, D. I. (1978). A case study of language awareness. In A. Sinclair, R. Jarvella, and W. J. M. Levelt (eds), *The Child's Conception of Language*. Berlin: SpringerVerlag, 45–64.
- Smet, H. De and Heyvaert, L. (2011). The meaning of the English present participle. *English Language and Linguistics* 15(3): 473–498.
- Smith, C. (1970). Jespersen's 'move and change' class and causative verbs in English. In M. A. Jazayery, E. C. Polomé, and W. Winter (eds), *Linguistic and Literary Studies in Honor of Archibald A. Hill*, vol. 2: *Descriptive Linguistics*. The Hague: Mouton, 101–109.

- Smith, C. (2000). The semantics of the Navajo verb base. In T. B. Fernald and P. R. Platero (eds), *The Athabaskan Languages. Perspectives on a Native American Language Family*. Oxford: Oxford University Press, 200–227.
- Smith, K. D. (1969). Sedang affixation. *Mon-Khmer Studies* 3: 108–124.
- Smith, L. R. (1978a). *A Survey of the Derivational Postbases of Labrador Inuttut (Eskimo)*. (National Museum of Man Mercury Series No. 45). Ottawa: Museum of Civilization.
- Smith, L. R. (1978b). Some properties of Labrador Inuttut verbal derivation. *Études/Inuit/Studies* 2(2): 37–48.
- Smith, L. R. (1980). Some categories and processes of Labrador Inuttut word formation. *International Journal of American Linguistics* 46(4): 279–288.
- Smith, R. L. (1973). Reduplication in Ngeq. *Mon-Khmer Studies* 4: 85–111.
- Smolka, E., Komlosi, S., and Rösler, F. (2009). When semantics means less than morphology: the processing of german prefixed verbs. *Language and Cognitive Processes* 24(3): 337–375.
- Snellgrove, D. (1954). Review of Marcelle Lalou's *Manuel élémentaire de tibétain classique*. *Bulletin of the School of Oriental and African Studies* 16(1): 198–199.
- Söderbergh, R. (1979). Barnets språkliga kreativitet belyst av exempel från ordböjning och ordbildning [The linguistic creativity of children: word-inflection and word-formation]. *Språkform och Språknorm* 67: 236–245.
- Solms, H.-J. (1990). Das System der Präfixverben in der frühesten Überlieferung des Hartmannschen 'Gregorius' (Hs. A aus dem Alemannischen des 13. Jahrhundert). In W. Besch (ed.), *Deutsche Sprachgeschichte. Grundlagen, Methoden, Perspektiven; Festschrift für Johannes Erben zum 65. Geburtstag*. Frankfurt am Main: Lang, 115–128.
- Solms, H.-J. (1998). Historische Wortbildung. In W. Besch (ed.), *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. (Handbücher zur Sprach- und Kommunikationswissenschaft, 2). 2nd edn. Berlin: de Gruyter, 596–610.
- Solomyak, O. and Marantz, A. (2009). Lexical access in early stages of visual word processing: A single-trial correlational MEG study of heteronym recognition. *Brain and Language* 108: 191–196.
- Solomyak, O. and Marantz, A. (2010). Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience* 22(9): 2042–2057.
- Song, J. J. (2011). *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press.
- Spalding, A. (1993). *Inuktitut: A Grammar of N. Baffin Dialects Vol. 2*. Winnipeg: Wuerz Publishing.
- Spalding, A. (1998). *Inuktitut: a Multi-dialectal Outline Dictionary (with an Aivilingmiutaq base)*. Iqaluit, NT: Nunavut Arctic College.
- Spencer, A. (1991). *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Cambridge: Cambridge University Press.
- Spencer, A. (1998). Morphophonological operations. In A. Spencer and A. Zwicky (eds), *The Handbook of Morphology*. Oxford: Blackwell, 123–143.
- Spencer, A. (2005). Towards a typology of 'mixed categories'. In C. O. Orgun and P. Sells (eds), *Morphology and the Web of Grammar. Essays in Memory of Steven G. Lapointe*. Stanford, CA: CSLI Publications, 95–138.
- Spencer, A. (in press). *Lexical Relatedness: A Paradigm-based Model*. Oxford: Oxford University Press.
- Spencer, A. and Zwicky, A. M. (eds) (1998). *The Handbook of Morphology*. Oxford: Blackwell.
- Spieler, D. H. and Balota, D. A. (1998). Naming latencies for 2820 words. [On-line], Available: <<http://www.artsci.wustl.edu/~dbalota/naming.html>>.

- Splett, J. (1993). *Althochdeutsches Wörterbuch: Analyse der Wortfamilienstrukturen des Althochdeutschen, zugleich Grundlegung einer zukünftigen Strukturgeschichte des deutschen Wortschatzes*. Berlin: de Gruyter.
- Spreng, B. (2012). Viewpoint aspect in Inuktitut: the syntax and semantics of antipassives. Ph.D. thesis. University of Toronto.
- Sproat, R. (1985). On deriving the lexicon. Ph.D. dissertation. MIT.
- Sproat, R. (1992). *Morphology and Computation*. Cambridge, MA: MIT Press.
- Sproat, R. (2005). Current morphological theory. In *SKASE Journal of Theoretical Linguistics* 2(2): 63–75.
- Stankiewicz, E. (1968). *Declension and Gradation of Russian Substantives*. The Hague: Mouton.
- Starosta, S. (1992). Sora combining forms and pseudo-compounding. *Mon-Khmer Studies* 18–19: 77–105.
- Starostin, S. (1986). Problema geneticheskoy obshchnosti altajskix jazykov [The problem of the genetic relationship of Altaic languages]. *PIAC* 29: 104–112.
- Starostin, S. (1991). *Altajskaja problema i proisxozhdenie japonskogo jazyka* [The Altaic problem and the origin of the Japanese language]. Moskva: Nauka.
- Starostin, S. (2000). The phonological system of Proto-Altaic. In *Problemy izuchenija dal'nego rodstva jazykov na rubezhe tret'ego tysjacheletija*. Moskva: Rossijskij gosudarstvennyj gumanitarnyj univerzitet, 222–224.
- Starostin, S. (2005). Response to Stefan Georg's review of the *Etymological Dictionary of the Altaic Languages*. *Diachronica* 22(2): 451–454.
- Starostin, S., Dybo, A. and Mudrak, O. (2003). *Etymological Dictionary of the Altaic Languages*, 3 vols. Leiden: Brill Academic Publishers.
- Statistics Canada (2006). <<http://www12.statcan.gc.ca/census-recensement/2006/dp-pd/prof/92-594/details/page.cfm?Lang=EandGeo1=CSDandCode1=6205015andGeo2=PRandCode2=01andData=CountandSearchText=ArviatandSearchType=ContainsandSearchPR=01andB1=AllandCustom=>>> accessed August 8, 2012.
- Steele, J. (2005). *The Aboriginal language of Sydney: A partial reconstruction of the indigenous language of Sydney based on the notebooks of William Dawes of 1790-91, informed by other records of the Sydney and surrounding languages to c.1905*, Master of Arts (Research), Macquarie University.
- Steele, S. (1995). Towards a theory of morphological information. *Language* 71: 260–309.
- Steels, L. (1997). The synthetic modeling of language origins. In H. Gouzoules (ed.), *Evolution of Communication*, vol. 1. Amsterdam: Benjamins, 1–34.
- Štekauer, P. (1996). *A Theory of Conversion in English*. Frankfurt am Main: Lang.
- Štekauer, P. (1998). *An Onomasiological Theory of English Word-Formation*. Amsterdam: Benjamins.
- Štekauer, P. (2005). Onomasiological approach to word-formation. In P. Štekauer and R. Lieber (eds), *Handbook of Word-formation*. Dordrecht: Springer, 207–232.
- Štekauer, P. (2006). Onomasiological theory of word formation. In K. Brown (ed.), *Encyclopedia of Language and Linguistics*. Oxford: Elsevier, 34–37.
- Štekauer, P. (2008). Conversion: a cross-linguistic research. Paper presented at the 13th Morphology Meeting. Vienna: University of Vienna.
- Štekauer, P. (2011). On some issues of diminutives from a cross-linguistic perspective. In S. D. Schmid, P. Gévaudan, W. Mihatsch, and W. Waltereit (eds), *Rahmen des Sprechens*. Tübingen: Narr, 425–435.

- Štekauer, P. (forthcoming a). The delimitation of derivation and inflection. In P. O. Müller, I. Ohnheiser, S. Olsen, F. Rainer (eds), *HSK Word-Formation. An International Handbook of the Languages of Europe*. Berlin: Mouton De Gruyter.
- Štekauer, P. (forthcoming b). Backformation. In P. O. Müller, I. Ohnheiser, S. Olsen, and F. Rainer, (eds), *HSK Word-formation. An International Handbook of the Languages of Europe*. Berlin: Mouton De Gruyter.
- Štekauer, P. and Lieber, R. (eds) (2005). *Handbook of Word-Formation*. Dordrecht: Springer.
- Štekauer, P. and Valera, S. (2007). *Universals, tendencies and typology in word-formation: a cross-linguistic research*. Manuscript.
- Štekauer, P., Chapman, D., Tomaščíková, S., and Franko, Š. (2005). Word-formation as creativity within productivity constraints. sociolinguistic evidence. *Onomasiology Online* (6): 1–55.
- Štekauer, P., Valera, S., and Körtvélyessy, L. (2012). *Word-Formation in the World's Languages. A Typological Survey*. Cambridge: Cambridge University Press.
- Stell, N. (1987). Gramática descriptiva de la lengua nivacle (chulupí). Ph.D. dissertation. University of Buenos Aires.
- Stephany, U. (1997). Diminutives in early Greek, a preliminary investigation. In W. Dressler (ed.), *Studies in Pre- and Protomorphology*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften, 147–156.
- Steriade, D. (1988). Reduplication and syllable transfer in Sanskrit. *Phonology* 5: 73–155.
- Sterling, C. (1982). The psychological productivity of inflectional and derivational morphemes. In D. Rogers, and J. Slobada (eds), *The Acquisition of Symbolic Skills*. New York: Plenum, 179–185.
- Stern, C. and Stern, W. (1928). *Die Kindersprache: Eine Psychologische und Sprachtheoretische Untersuchung*. 4th rev. edn. Leipzig: Barth.
- Stevens, A. M. (1997). Truncation phenomena in contemporary Indonesian. In C. Odé and W. Stokhof (eds), *Proceedings of the Seventh International Conference on Austronesian Linguistics*. Amsterdam: Rodopi, 167–181.
- Stockwell, R. and Minkova, D. (2001). *English Words: History and Structure*. Cambridge: Cambridge University Press.
- Stolz, C. and Stolz, T. (2001). Mesoamerica as a linguistic area. In M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds), *Language Typology and Language Universals: An International Handbook*. Vol. 2 (Handbücher zur Sprach- und Kommunikationswissenschaft, 20.2). New York: Walter de Gruyter, 1542–1553.
- Stolz, T. (1991). *Sprachbund im Baltikum? Estnisch und Lettisch im Zentrum einer sprachlichen Konvergenzlandschaft* (Bochum-Essener Beiträge zur Sprachwandelforschung, 13). Bochum: N. Brockmeyer.
- Stolz, T., Stroh, C. and Urdze, A. (2011). *Total Reduplication: The Areal Linguistics of a Potential Universal*. Berlin: Akademie-Verlag.
- Stonham, J. (1994). *Combinatorial Morphology*. Amsterdam: Benjamins.
- Storch, A. and Vossen, R. (2007). Odours and colours in Nilotic: comparative case studies. In D. Payne and M. Reh (eds), *Advances in Nilo-Saharan Linguistics*. Cologne: Rüdiger Köppe, 223–240.
- Stowell, T. (1991). The alignment of arguments in adjective phrases. In S. Rothstein (ed.), *Perspectives on Phrase Structure: Heads and Licensing. Syntax and Semantics* 25. San Diego: Academic Press, 105–135.
- Street, J. C. (1962). Review of N. Poppe, *Vergleichende Grammatik der altaischen Sprachen, Teil I* (1960). *Language* 38: 92–98.

- Stump, G. T. (1991). A paradigm-based theory of morphosemantic mismatches. *Language* 67: 675–725.
- Stump, G. (1992). Position classes and morphological theory. In G. Booij and J. van Marle (eds), *Yearbook of Morphology 1992*. Dordrecht: Kluwer, 129–180.
- Stump, G. T. (1993). How peculiar is evaluative morphology? *Journal of Linguistics* 29: 1–36.
- Stump, G. T. (1998). Inflection. In A. Spencer and A. Zwicky (eds), *Handbook of Morphology*. Oxford: Blackwell, 13–43.
- Stump, G. (2001). *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Stump, G. T. (2005). Word-formation vs inflectional morphology. In R. Lieber and P. Štekauer (eds), *Handbook of English Word-formation*. Dordrecht: Springer, 49–71.
- Stump, G. T. (2010). The derivation of compound ordinal numerals: Implications for morphological theory. *Word Structure* 3(2): 205–233.
- Sugioka, Y. and Lehr, R. (1983). Adverbial *-ly* as an inflectional affix. *Chicago Linguistic Society* 19(2): 293–300.
- Sun, J. T.-S. (2000a). Parallelisms in the verb morphology of Sidaba rGyalrong and Guanyinqiao in rGyalrongic. *Language and Linguistics* 1(1): 161–190.
- Sun, J. T.-S. (2000b). Stem alternations in Puxi verb inflection. *Language and Linguistics* 1(2): 211–232.
- Sun, J. T.-S. (2003). Caodeng rGyalrong. In R. LaPolla and G. Thurgood (eds), *The Sino-Tibetan Languages*. London: Routledge, 490–502.
- Sun, J. T.-S. (2004a). 草登嘉戎語的狀貌詞 Caodeng Jiarongyu de zhuangmaoci [Ideophones in Caodeng Rgyalrong]. *Minzu Yuwen* 5: 1–11.
- Sun, J. T.-S. (2004b). Verb-stem variations in Showu rGyalrong. In Y.-C. Lin et al. (eds), *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-Cherng Gong on His Seventieth Birthday*. Taipei: Institute of Linguistics, Academia Sinica, 269–296.
- Sun, J. T.-S. (2012). Complementation in Caodeng rGyalrong. *Language and Linguistics* 13(3): 471–498.
- Sun, J. T.-S. and Shidanluo (2002). 草登話與認同等第相關的語法現象 Caodenghua yu rentong dengdi xiangguan de yufa xianxiang [Empathy Hierarchy in Caodeng Rgyalrong Grammar]. *Language and Linguistics* 3(1): 79–99.
- Sundby, B. (1995). *English Word-formation as Described by English Grammarians 1600–1800* (Studia Anglistica Norvegica 7). Oslo: Novus.
- Sunik, O. (1962). *Glagol v tunguso-manchzhurskix jazykax* [The verb in the Tungus-Manchu languages]. Moscow: Akademija Nauk.
- Sunik, O. (1968). Udegejskij jazyk [The Udihe language]. In V. Vinogradov (ed.), *Jazyki narodov SSSR. Vol. 5. Mongol'skie, tunguso-man'chzhurskie i paleoaziatskie jazyki*. Leningrad: Akademija Nauk, 210–232.
- Sunik, O. (1997). Udegejskij jazyk [The Udihe language]. In V. Jarceva (ed.), *Jazyki mira. Mongol'skie jazyki. Tunguso-Manchzhurskie jazyki. Japonskij jazyk. Korejskij jazyk*. Moscow: Indrik, 236–248.
- Surabhi, B. (1994). *Aspects of the Phonology of Hindi and English*. New Delhi: Arnold Publishers.
- Sutton, P. (1995). *Wik-Ngathan Dictionary*. Adelaide: Caitlin Press.
- Suwilai, P. (2011). Pearic, a dying branch of Austroasiatic languages and its struggle for survival. In S. Srichampa and P. Sidwell (ed.), *Austroasiatic Studies: Papers from ICAAL 4, Mon-Khmer Studies Journal Special Issue* 2: 138–153.

- Suzuki, R. (1999). Language socialization through morphology: the affective suffix -chau in Japanese. *Journal of Pragmatics* 31: 1423–1441.
- Suzuki, S. (1990). On the Origin and Development of the Action Noun Forming Suffix -nis in Old English. *Indogermanische Forschungen: Zeitschrift für Indogermanistik und Allgemeine Sprachwissenschaft* 95: 184–207.
- Svaib, T. (1993). Are doggies really little? Evidence from children's use of diminutives. In E. V. Clark (ed.), *Proceedings of the 24th annual Child Language Research Forum*. Stanford, CA: CSLI, 149–161.
- Svantesson, J.-O. (1983). *Kammu Phonology and Morphology* (Travaux de L'Institute de Linguistique de Lund XVIII). Sweden: CWK Gleerup.
- Svenonius, P. (2008). Projections of P. In A. Asbury, J. Dotlacil, B. Gehrke, and R. Nouwen (eds), *Syntax and Semantics of Spatial P*. Amsterdam: Benjamins, 63–84.
- Swadesh, M. (1959). *Mapas de clasificación lingüística de México y las Américas*. México: Universidad Nacional Autónoma de México.
- Swan, O. (2002). *A Grammar of Contemporary Polish*. Bloomington, IN: Slavica Publishers.
- Sweet, H. (1891–1898). *A New English Grammar: Logical and Historical*. Oxford: Clarendon Press.
- Szymanek, B. (1988). *Categories and Categorization in Morphology*. Lublin: Redakcja Wydawnictw Katolickiego Uniwersytetu Lubelskiego.
- Szymanek, B. (2009). IE, Slavonic: Polish. In R. Lieber and P. Štekauer (eds), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 464–477.
- Szymanek, B. (2010). *A Panorama of Polish Word-Formation*. Wydawnictwo KUL: Lublin.
- Szymanek, B. and Derkach, T. (2005). Constraints on the derivation of double diminutives in Polish and Ukrainian. *Studies in Polish Linguistics* 2: 93–112.
- Tạ, Văn Thông (1998). Phương thức Láy trong Tiếng Koho (Methods of Reduplication in Koho). *Tứ Láy: Những Vấn đề còn Để Ngó* (Reduplication: Issues to Consider). Hanoi: Nhà Xuất Bản Khoa Học Xã Hội, 194–213.
- Tabak, W., Schreuder, R., and Baayen, R. H. (2005). Lexical statistics and lexical processing: semantic density, information complexity, sex, and irregularity in Dutch. In S. Kepser and M. Reis (eds), *Linguistic Evidence—Empirical, Theoretical, and Computational Perspectives*. Berlin: Mouton de Gruyter, 529–555.
- Tabak, W., Schreuder, R., and Baayen, R. H. (2010). Producing inflected verbs: a picture naming study. *The Mental Lexicon* 5(1): 22–46.
- Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning and Verbal Behavior* 20: 289–297.
- Taft, M. (1991). *Reading and the Mental Lexicon*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes* 9: 271–294.
- Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology* 57A: 745–765.
- Taft, M. and Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14: 638–647.
- Taft, M. and Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior* 15: 607–620.
- Taine-Cheikh, C. (2002). Morphologie et morphogenèse des diminutifs en zénaga (berbère de Mauritanie). In K. Naït-Zerrad (ed.), *Articles de linguistique berbère. Mémorial Werner Vycichl*. Paris: l'Harmattan, 427–454.

- Taine-Cheikh, C. (2003). L'adjectif et la conjugaison suffixale en berbère. In J. Lentin and A. Roth (eds), *Mélanges David Cohen*. Paris: Maisonneuve et Larose, 661–674.
- Talamo, L. (2011). Complexity-Based Ordering in Italian: some new data. Paper presented at the 3rd Vienna Workshop on Affix Order: Advances in Affix Order Research. Vienna, January 15–16.
- Tarp, S. (2009). Homonymy and polysemy in a lexicographical perspective. *Zeitschrift für Anglistik und Amerikanistik* 57: 289–305.
- Taylor, J. R. (1992). *Linguistic Categorization*. Orford: Oxford University Press.
- Teichmann, C. G. and Schürmann, C. W. (1840). *Outlines of a grammar, vocabulary, and phraseology, of the aboriginal language of South Australia, spoken by the natives in and for some distance around Adelaide*. Adelaide: Published by the authors, at the native location.
- Teleman, U. (2005). The standard languages and their systems in the 20th century IV: Swedish. In O. Bandle, K. Braunmüller, E. H. Jahr, A. Kärker, and H.-P. Naumann (eds), *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, vol. 2. Berlin: Walter de Gruyter, 1603–1626.
- Temmerman, R. (2000). *Towards New Ways of Terminology Description: The Sociocognitive Approach*. Amsterdam: Benjamins.
- ten Hacken, P. (1994). *Defining Morphology: A Principled Approach to Determining the Boundaries of Compounding, Derivation, and Inflection*. Hildesheim: Olms.
- ten Hacken, P. (2009). Word Manager. In C. Mahlow and M. Piotrowski (eds), *State of the Art in Computational Morphology*. Berlin: Springer, 88–107.
- ten Hacken, P. (2010a). Creating Legal Terms: A Linguistic Perspective. *International Journal for the Semiotics of the Law* 23: 407–425.
- ten Hacken, P. (2010b). The tension between definition and reality in terminology. In A. Dykstra and T. Schoonheim (eds), *Proceedings of the XIV Euralex International Congress*. Ljouwert: Fryske Akademy/Afuk, 915–927.
- ten Hacken, P. (2010c). Synthetic and exocentric compounds in a parallel architecture. In S. Olsen (ed.), *New Impulses in Word-Formation*. Linguistische Berichte Sonderheft 17: 233–251.
- Tenishev, E. R. (1976). *Stroj salarskogo jazyka* [The structure of Salar]. Moscow: Nauka.
- Terasawa, J. (2001). The scarcity of formations in -ere in Old English poetry. *Anglia: Zeitschrift für Englische Philologie* 119(2): 193–206.
- Tersis, N. (2008). *Forme et sens des mots du Tunumiisut. Lexique inuit du Groenland oriental*. Paris: Peters.
- Thieberger, N. (2006). *A Grammar of South Efate: An Oceanic Language of Vanuatu*. Oceanic Linguistics Special Publication No. 33. Honolulu: University of Hawai'i Press.
- Thomas, D. M. (1969). Chrau affixes. *Mon-Khmer Studies* 3: 90–107.
- Thomas, D. (1971). *Chrau Grammar* (Oceanic Linguistics Special Publication 7). Honolulu: University of Hawaii Press.
- Thomas, D. (1987–8). The instrument/locative and goal affix -N- in Surin Khmer. *Mon-Khmer Studies* 16–17: 85–98.
- Thomas, D. M. (1990). The instrument/locative and goal affix -N- in Surin Khmer. *Mon-Khmer Studies* 16–17: 85–98.
- Thomason, S. G. (2001). *Language Contact*. Edinburgh: Edinburgh University Press.
- Thomason, S. G. and Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Thompson, C. (1993). The Areal Prefix *hu-* in Koyukon Athapaskan. *International Journal of American Linguistics* 59: 315–333.

- Thompson, L. (1965). *A Vietnamese Grammar*. Seattle: University of Washington Press.
- Thompson, P. (2012). The lexicalisation of syncope: derivational affixes in West Saxon adjectives. In H. Sauer and G. Waxenberger (eds), *English Historical Linguistics 2008. Selected papers from the fifteenth International Conference on English Historical Linguistics (ICEHL 15), Munich, 24–30 August 2008. Volume II: Words, texts and genres*. Amsterdam: Benjamins, 133–146.
- Thráinsson, H. (1994). Icelandic. In E. König and J. van der Auwera (eds), *The Germanic Languages*. London: Routledge, 142–189.
- Thurgood, G. (1999). *From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change: With an Appendix of Chamic Reconstructions and Loanwords* (Oceanic Linguistics Special Publications, No. 28). Pacific Linguistics.
- Thurgood, G. and LaPolla, R. J. (2003). *The Sino-Tibetan Languages*. London: Routledge.
- Till, W. C. (1970). *Koptische Grammatik*. 4th edn. Leipzig: Verlag Enzyklopädie.
- Timberlake, A. (2004). *A Reference Grammar of Russian*. Cambridge: Cambridge University Press.
- Tinbergen, D. C. (1919). Kinderpraat. *De Nieuwe Taalgids* 13(1–16): 65–86.
- Tompá, J. (ed.) (1961). *A mai magyar nyelv rendszere, Leíró nyelvtan I.–II.* [The linguistic system of contemporary Hungarian. Descriptive Grammar Vols I.–II.]. Budapest: Akadémiai Kiadó.
- Toosarvandani, M. (2010). Patterns of nominalization in Numic. *International Journal of American Linguistics* 76(1): 71–100.
- Topping, D. M. (1973). *Chamorro Reference Grammar*. Honolulu: University Press of Hawaii.
- Tournadre, N. (1996). *L'ergativité en tibétain. Approche morphosyntaxique de la langue parlée*. Louvain: Peeters.
- Tournadre, N. (2008). Arguments against the Concept of 'Conjunct'/'Disjunct' in Tibetan. In B. Huber et al. (eds), *Chomolangma, Demawend und Kasbek. Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*. Vol 1. Halle (Saale): International Institut for Tibetan and Buddhist Studies, 281–308.
- Tournier, J. (1985). *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Genève: Slatkine.
- Torre Alonso, R. (2011). Affix combination in Old English noun formation: distribution and constraints. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics* 24: 257–278.
- Tovar, A. (1951). Un capítulo de lingüística general. *Boletín de la Academia Argentina de Letras*, Tomo XX, No. 77 Buenos Aires: Recta Sustenta, 369–403.
- Tovar, A. (1961). *Catálogo de Lenguas de América del Sur*. Buenos Aires: Sudamericana.
- Tovar, A. (1964). El grupo mataco y su relación con otras lenguas de América del Sur. *Actas del 35° Congreso Internacional de Americanistas*. México: Tomo II, 439–452.
- Toweett, T. (1979). *A Study of Kalenjin Linguistics*. Nairobi: Kenya Literature Bureau.
- Tremblay, A. and Baayen, R. H. (2010). Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (ed.), *Perspectives on Formulaic Language: Acquisition and Communication*. London: Continuum International Publishing Group, 151–173.
- Tremblay, A. and Tucker, B. V. (2011). The effects of N-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon* 6(2): 302–324.

- Tremblay, A., Derwing, B., Libben, G., and Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning* 61(2): 569–613.
- Trips, C. (2007). Lexical semantics and diachronic morphology. The development of -hood, -dom and -ship in the history of English. Habilitationsschrift. Stuttgart: Universität Stuttgart.
- Trips, C. (2009). *Lexical Semantics and Diachronic Morphology. The Development of -hood, -dom and -ship in the History of English*. Tübingen: Niemeyer.
- Trips, C., and Stein, A. (2008). 'Was French -able borrowable? a diachronic study of word-formation processes due to language contact'. In R. Dury, M. Gotti, and M. Dossena (eds), *English Historical Linguistics 2006. Selected Papers from the Fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, August, 21–25 2006*. 2. *Lexical and semantic change*, 3 vols, 296. Amsterdam: Benjamins, 217–41.
- Trnka, B. (1969). Conversion in English. *Brno Studies in English* 8: 183–187.
- Trnková, J. (1995). O substantívnej deminutívnej derivácii v slovenčine. *Slovenská reč* 56(2): 87–95.
- Trommer, J. (2012). Ø-exponence. In J. Trommer (ed.), *The Morphology and Phonology of Exponence*. Oxford: Oxford University Press, 326–354.
- Trommer, J. and Zimmermann, E. (2010). Generalized mora affixation. Paper presented at the 28th meeting of the West Coast Conference on Formal Linguistics, University of Southern California, Los Angeles, CA, February 21.
- Trubetzkoy, N. (1934). Das morphologische System der russischen Sprache. *Travaux du Cercle Linguistique de Prague*, Vol. 5.
- Tsujimura, N. (2007). *An Introduction to Japanese Linguistics*. 2nd edn. Oxford: Blackwell.
- Tsujimura, N. and Davis, S. (2011a). A construction approach to innovative verbs in Japanese. *Cognitive Linguistics* 22: 797–823.
- Tsujimura, N. and Davis, S. (2011b). Construction morphology in Japanese and implications to prosodic morphology. MS, Indiana University
- Tucker, A. N. (1967). *The Eastern Sudanic Languages*. London: Dawsons of Pall Mall for the International African Institute.
- Tucker, A. N. and Bryan, M. A. (1966). *Linguistic Analyses: The Non-Bantu Languages of Northeastern Africa*. London: Oxford University Press for the International African Institute.
- Tucker, A. N. and Tompo ole Mpaayei, J. (1955). *A Maasai Grammar with Vocabulary*. London: Longmans, Green and Co.
- Tuggy, D. (2005). Cognitive approach to word formation. In P. Štekauer and R. Lieber (eds), *Handbook of Word Formation*. Dordrecht: Springer, 233–266.
- Tuttle, S. and Hargus, S. (2004). Explaining variability in affix order: the Athabaskan areal and third person prefixes. *Alaska Native Language Center Working Papers* 4: 70–98.
- Tyler, A. and Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language* 28: 649–667.
- Ubrjatova, E. I. et al. (eds) (1982). *Grammatika sovremennogo jakutskogo literaturnogo jazyka. Fonetika i morfologija* [A grammar of modern literary Sakha: Phonetics and morphology]. Moscow: Nauka.
- Uebach, H. and Zeisler, B. (2008). rje-blas, pha-los and Other Compounds with Suffix -s. In B. Huber, M. Volkart, and P. Widmer (eds), *Old Tibetan Texts. Chomolangma, Demawend und Kasbek. Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*, vol 1. Halle: International Institute for Tibetan and Buddhist Studies, 309–334.
- Uhlenbeck, E. M. (1962). Limitations of morphological processes: some preliminary remarks. *Lingua* 11: 426–432.

- Uhlenbeck, E. M. (1978). *Studies in Javanese Morphology*. The Hague: Nijhoff.
- Ullman, M. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92: 231–270.
- Ullman, M. T., Estabrooke, I. V., Steinhauer, K., Brovett, C., Pancheva, R., Ozawa, K., Mordecai, K., and Maki, P. (2002). Sex differences in the neurocognition of language. *Brain and Language* 83: 141–143.
- Ultan, R. (1975). Infixes and their origins. In H. Seiler (ed.), *Linguistic workshop III. Arbeiten des Kölner Universalienprojekts 1974*. München: Fink (Structura, 9), 157–205.
- Ultan, R. (1978). Size-sound symbolism. In J. Greenberg (ed.), *Universals of Human Language. Vol 2*. Stanford, CA: Stanford University Press, 525–568.
- Unseth, P. (1998). Two old causative affixes in Surmic languages. In G. J. Dimmendaal and M. Last (eds), *Surmic Languages and Cultures*. Cologne: Rüdiger Köppe, 113–126.
- Uray, G. (1952). A Tibetan diminutive suffix. *Acta Orientalia Academiae Scientiarum Hungaricae* 2: 182–220.
- Uray, G. (1953a). Some problems of the ancient Tibetan verbal morphology: methodological observations on recent studies. *Acta Linguistica Academiae Scientiarum Hungaricae*, 37–62.
- Uray, G. (1953b). The suffix -e in Tibetan. *Acta Orientalia Academiae Scientiarum Hungaricae* 3: 229–244.
- Uray, G. (1955). Duplication, gemination and triplication in Tibetan. *Acta Orientalia Academiae Scientiarum Hungaricae* 4(1–3): 177–256.
- Urbanczyk, S. (2006). Reduplicative form and the root-affix asymmetry, *Natural Language and Linguistic Theory* 24: 179–240.
- Ushakova, T. N. (1969). O mekhanizmaxh detskogo slovotvorchestva [On mechanisms in children's word creation]. *Voprosy Psikhologii* 15: 62–74.
- Ushakova, T. N. (1976). Children's word creation. In J. Průcha (ed.), *Soviet Studies in Language and Language Behavior*. Amsterdam: North-Holland, 165–175.
- Ussishkin, A. (1999). The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-to-output correspondence. *Phonology* 16: 401–442.
- Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language and Linguistic Theory* 23(1): 169–218.
- Ussishkin, A. (2006). Affix-favored contrast inequity and psycholinguistic grounding for nonconcatenative morphology. *Morphology* 16(1): 107–125.
- Ussishkin, A. (2011). Tier segregation. In M. van Oostendorp, C. Ewen, E. Hume, and K. Rice (eds), *The Blackwell Companion to Phonology*. Oxford: Wiley Blackwell, 1838–1865.
- Vajda, E. J. (2004). *Ket* (Languages of the world/materials 204). Munich: Lincom.
- Vajda, E. J. (2007). Ket morphology. In A. Kaye (ed.), *Morphologies of Asia and Africa*. Winona Lake, IN: Eisenbrauns, 1277–1325.
- Vajda, E. J. (2008). Head-negating enclitics in Ket. In E. Vajda (ed.), *Subordination and Coordination Strategies in North Asian Languages*. Amsterdam: Benjamins, 179–201.
- Vajda, E. J. (2009). Loanwords in Ket. In M. Haspelmath and U. Tadmor (eds), *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: Mouton de Gruyter, 471–494.
- Vajda, E. J. (2011). A Siberian link to Na-Dene languages. In J. Kari and B. Potter (eds), *The Dene Yeniseian Connection*. Fairbanks, AK: ANLC, 33–99.
- Vajda, E. J. (2013). Metathesis and reanalysis in Ket. *Tomsk Journal of Linguistics and Anthropology* 1(1): 14–26.
- van den Berg, R. (1989). *A Grammar of the Muna Language*. Dordrecht: Foris.

- van der Lugt, A. (1999). *From Speech to Words*. Nijmegen: Catholic University.
- van der Voort, H. (2009). Reduplication and repetition of person markers in Guaporé isolates. *Morphology* 19: 263–286.
- van Driem, G. (1987). *A Grammar of Limbu*. Berlin: Mouton.
- van Driem, G. (2003). Tibeto-Burman vs. Sino-Tibetan. In W. Winter, B. L. M. Bauer, and G.-J. Pinault (eds), *Language in Time and Space: A Festschrift for Werner Winter on the occasion of his 80th birthday*. Berlin: Walter de Gruyter, 101–119.
- van Driem, G. (2011). The Trans-Himalayan phylum and its implications for population prehistory. *Communication on Contemporary Anthropology* 5: 135–142. [Available at: <<http://www.himalayanlanguages.org/files/driem/pdfs/2011The%20Trans-Himalayan%20phylum%20and%20its%20implications%20for%20population%20prehistory.pdf>>. accessed August 15, 2012.]
- van Driem, G. (2012). The Trans-Himalayan phylum and its implications for population prehistory. *Communication on Contemporary Anthropology* 5: 135–142.
- van Geenhoven, V. (1998a). *Semantic Incorporation and Indefinite Descriptions: Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. Dissertations in Linguistics, Stanford: CSLI Publications.
- van Geenhoven, V. (1998b). On the argument structure of some noun incorporating verbs in West Greenlandic. In M. Butt and W. Geuder (eds), *The Projection of Arguments: Lexical and Compositional Factors*. Stanford, CA: CSLI, 225–263.
- van Hout, A. and Roeper, T. (1998). Events and aspectual structure in derivational morphology. *MIT Working Papers in Linguistics* 32: 175–220.
- van Hout, R. and Muysken, P. (1994). Modeling lexical borrowability. *Language Variation and Change* 6: 39–62.
- van Marle, J. (1985). *On the Paradigmatic Dimension of Morphological Creativity*. Dordrecht: Foris.
- van Marle, J. (1986). The Domain Hypothesis: The study of rival morphological processes. *Linguistics*, 24: 601–627.
- van Marle, J. (1994). Paradigms. In R. E. Asher and J. Simpson (eds), *Encyclopedia of Language and Linguistics*. Oxford: Pergamon, 6: 2927–2930.
- van Valin, R. D. and LaPolla, R. J. (1997). *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.
- van Voorst, J. (1996). Some systematic differences between the Dutch, French, and English transitive construction. *Language Sciences* 18: 227–245.
- Vannest, J., Newport, E. L., Newman, A. J., and Bavelier, D. (2011). Interplay between morphology and frequency in lexical access: The case of the base frequency effect. *Brain Research* 1373: 144–159.
- Vasilevič, G. (1960). K voprosu o klassifikacii tunguso-man'chzhurskix jazykov [On the classification of the Tungus-Manchu languages]. *Voprosy Jazykoznanija* 2: 43–49.
- Vaux, B. (2003). Syllabification in Armenian, Universal Grammar, and the lexicon. *Linguistic Inquiry* 34: 91–125.
- Vázquez-López, V. (2012). Nominalisations in Early Modern English: internal structure, development and suffixal productivity. In D. Tizón-Couto (ed.), *New Trends and Methodologies in Applied English Language Research II. Studies in language variation, meaning and learning; [a selection of papers presented at the Second ELC International Postgraduate Conference on English Linguistics (ELC2), held at the University of Vigo in October 2009]*. Bern: Lang, 223–46.
- Velupellai, V. (2012). *An Introduction to Linguistic Typology*. Amsterdam: Benjamins.

- Veneziano, E. and Clark, E. V. (2012). Acquiring verb constructions in French: the Left Edge hypothesis. Unpublished MS, Stanford University.
- Verma, M. K. (2003). Bhojpuri. In G. Cardona, and D. Jain (eds), *The Indo-Aryan Languages*. London: Routledge, 515–537.
- Verschik, A. (2004). Estonian compound nouns and their equivalents in the local variety of Russian. *Scando-Slavica* 50: 93–109.
- Veselinova, L. (2007). Suppletion from a typological perspective. In M. Miestamo and B. Wälchli (eds), *New Challenges in Typology*. Berlin: Mouton de Gruyter, 127–151.
- Vidal, A. (2001). Pilagá grammar (Guaykuruan family, Argentina). Ph.D. dissertation. University of Oregon.
- Vidal, A. and Nercesian, V. (2005a). Sustantivo y verbo en wichí. Hacia una taxonomía de clases de palabras. *Liames* 5: 7–25.
- Vidal, A. and Nercesian, V. (2005b). Causativos en wichí (mataco-mataguaya). *Proceedings from the II Congreso de Idiomas Indígenas de Latinoamérica*. Austin, TX.
- Vidal, A. and Nercesian, V. (2009a). Loanwords in Wichí, a Mataco-Mataguayan language of Argentina. In M. Haspelmath and U. Tadmor (eds), *Loanwords in the World's Languages. A Comparative Handbook of Loanword Typology*. The Hague: Mouton de Gruyter, 1015–1034.
- Vidal, A. and Nercesian, V. (2009b). Wichí vocabulary. In M. Haspelmath and U. Tadmor (eds), *World Loanword Database*. Munich: Max Planck Digital Library.
- Viegas Barros, P. (1993). ¿Existe una relación genética entre las lenguas mataguayas y guaycurúes? *Hacia una Nueva Carta Étnica del Gran Chaco* 5: 193–213.
- Viegas Barros, P. (2002). Fonología del Proto-Mataguayo. Las fricativas dorsales. In M. Crevels, S. van de Kerke, S. Meira, and H. van der Voort (eds), *Current Studies on South American Languages. Indigenous Languages of Latin America*, 3. Leiden: University of Leiden, 137–148.
- Viên Ngôn Ngữ Học (Institute of Linguistics). (1995). *Từ Điển Từ Láy Tiếng Việt* (A Dictionary of Reduplication in Vietnamese). Hanoi: Nhà Xuất Bản Giáo Dục.
- Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition* 28(4): 735–747.
- Voeltz, E. (1977). Proto Niger-Congo extensions. Ph.D. dissertation. University of California, Los Angeles.
- Volkart, M. (2003). Types of compounds in Written Tibetan. In T. R. Kansakar and M. Turin (eds), *Themes in Himalayan Languages*. Heidelberg: South Asia Institute and Kathmandu: Tribhuvan University, 232–248.
- von Fintel, K. and Matthewson, L. (2008). Universals in semantics. *The Linguistic Review* 25: 139–201.
- Vooy, C. G. N. de (1916). Iets over woordvorming en woordbeteekenis in kindertaal. *De Nieuwe Taalgids* 10: 93–100, 128–141.
- Vorbichler, A. (1965). *Die Phonologie und Morphologie des Balese (Ituri-Urwald, Kongo)*. Glückstadt: Augustin.
- Vovin, A. (1993). Towards a new classification of Tungusic languages. *Eurasian Studies Yearbook* 65: 99–113.
- Vovin, A. (2005). The end of the Altaic controversy (review of Starostin et al. 2003). *Central Asiatic Journal* 49(1): 71–132.

- Vu, S. (2007). A Unified analysis of some Vietnamese reduplication forms. In M. Alves, P. Sidwell, and D. Gil (eds), *SEALS VII: Papers from the 8th Annual Meeting of the Southeast Asian Linguistics Society 1998*, 165–192.
- Vul, E., Harris, C., Winkielman, P., and Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science* 4(3): 274–290.
- Vương, H. L. (1997). *A Bru-Vietnamese-English Dictionary*. Hue, Vietnam: Nhà Xuất Bản Thuận Hóa.
- Vycichl, W. (1988). Beja: A language with seven seals. In M. Bechhaus-Gerst and F. Servisko (eds), *Papers from the International Symposium on Cushitic and Omotic Languages*. Hamburg: Helmut Buske, 411–432.
- Vycichl, W. (2005). *Berberstudien and a sketch of Siwi Berber* (Dymitr Ibrizimow and Maarten Kossmann (eds). Köln: Rüdiger Köppe.
- Waag, C. (2010). *The Fur Verb and its Context*. Cologne: Rüdiger Köppe.
- Wade, T. (2002). *Russian Grammar and Verbs*. Oxford: Oxford University Press.
- Wagner, A. and Rescorla, R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black and W. F. Prokasy (eds), *Classical Conditioning II*. New York: Appleton-Century-Crofts, 64–99.
- Wagner-Nagy, B. B. (2001), *Die Wortbildung im Nganasanischen*. Szeged: Universitas Szegediensis/Benjamins.
- Walker, A. and Zorc, R. D. (1981). Austronesian loanwords in Yolngu-Matha of northeast Arnhem Land. *Aboriginal History* 5(2): 109–134.
- Walker, J. (1924). *The Rhyming Dictionary*, rev. L. Dawson. London: Routledge.
- Walker, R. (2011). Nasal harmony. In M. van Oostendorp, C. Ewen, E. Hume, and K. Rice (eds), *The Blackwell Companion to Phonology*. Oxford: Wiley Blackwell, 2516–2537.
- Wangka Maya Pilbara Aboriginal Language Centre (2010). *Nyangumarta e-Dictionary 2010* [online text], <<http://www.wangkamaya.org.au/index.php/pilbara-languages/languages-of-the-pilbara/nyangumarta-mainmenu-341/interactive-e-dictionary>>.
- Wasow, T. (1977). Transformations and the lexicon. In P. Culicover, T. Wasow, and A. Akmajian (eds), *Formal Syntax*. New York: Academic Press, 327–360.
- Waszakowa, K. (1993). *Słownictwo współczesnego języka polskiego. Rzeczowniki z formantami paradygmatycznymi*. Warsaw: Wydawnictwa Uniwersytetu Warszaw.
- Watson, J. (2006). Arabic morphology: diminutive verbs and diminutive nouns in San'ani Arabic. *Morphology* 16: 189–204.
- Watson, R. (1966a). Clause to sentence gradations in Pacoh. *Lingua* 16: 166–188.
- Watson, R. (1966b). *Reduplication in Pacoh*. Hartford, CT: Hartford Seminary Foundation thesis.
- Watson, R. (1976). Pacoh numerals. *Linguistics* 74: 81–88.
- Watson, R. (2011). A case for clitics in Pacoh. In S. Srichampa, P. Sidwell, and K. Gregerson (eds), *Austroasiatic Studies: Papers from ICAAL4. Mon-Khmer Studies*, Journal Special Issue No. 3. Dallas, SIL International; Salaya, Mahidol University; Canberra, Pacific Linguistics, 222–232.
- Watson, S. K. (1964). Personal pronouns in Pacoh. *Mon-Khmer Studies* 1: 81–97.
- Watson, S. (1966). Verbal affixation in Pacoh. *Mon-Khmer Studies* 2: 15–30.
- Watts, J. (2010). Ancient tribal language becomes extinct as last speaker dies, *The Guardian*, February 4. <<http://www.guardian.co.uk/world/2010/feb/04/ancient-language-extinct-speaker-dies>>.

- Weingarten, R., Nottbusch, G., and Will, U. (2004). Morphemes, syllables and graphemes in written word production. In T. Pechmann and C. Habel (eds), *Multidisciplinary Approaches to Speech Production*. Berlin: Mouton de Gruyter, 529–572.
- Weinreich, U. ([1953] 1964). *Languages in Contact*. The Hague: Mouton.
- Weiss, D. (2009). Phonologie et morphosyntaxe du maba. Ph.D. thesis. Université Lumière Lyon 2.
- Wells, R. (1947). Immediate constituents. *Language* 23(2): 81–117.
- Wermser, R. (1976). *Statistische Studien zur Entwicklung des englischen Wortschatzes*. Schweizer anglistische Arbeiten, 91. Bern: Francke.
- Werner, H. (1997a). *Abriss der kottischen Grammatik*. Wiesbaden: Harrassowitz.
- Werner, H. (1997b). *Das Jugische (Sym-Ketische)*. Wiesbaden: Harrassowitz.
- Werner, H. (1997c). *Die ketische Sprache*. Wiesbaden: Harrassowitz.
- Werner, H. (1998). *Probleme der Wortbildung in den Jenissej-Sprachen*. Wiesbaden: Harrassowitz.
- Werner, H. (2005). *Die Jenissej-Sprachen des 18. Jahrhunderts*. Wiesbaden: Harrassowitz.
- Wetzer, H. (1996). *The Typology of Adjectival Predication*. Berlin: Mouton de Gruyter.
- Whaley, L. J. (1997). *Introduction to Typology*. Thousand Oaks, CA: Sage.
- Wheeler, M. W. (1998). Catalan. In M. Harris and N. Vincent (eds), *The Romance Languages*. London: Croom Helm, 170–208.
- Whitney, W. D. (1881). On mixture in language. *Transactions of the American Philosophical Association* 12: 1–26.
- Whitney, W. D. (1889). *Sanskrit Grammar*. 2nd edn. Cambridge, MA: Harvard University Press.
- Whitney, W. D. (1996). *Sanskrit Grammar*. 3rd edn. Boston: Ginn and Co.
- Whorf, B. L. (1935). The comparative linguistics of Uto-Aztecan. *American Anthropologist* 37: 600–608.
- Whorf, B. L. (1937/1956). Grammatical categories. In J. B. Carroll (ed.), *Language, thought and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press, 87–101.
- Wiebusch, T. and Tadmor, U. (2009). Loanwords in Mandarin Chinese. In M. Haspelmath and U. Tadmor (eds), *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: Mouton de Gruyter, 575–598.
- Wiedenhof, J. (2004). *Grammatica van het Mandarijn*. Amsterdam: Bulaaq.
- Wierzbicka, A. (1980). *Lingua mentalis: the semantics of natural language*. Sydney: Academic.
- Wierzbicka, A. (1989). The semantic primitives: the expanding set. *Quaderni di Semantica* 10(2): 309–332.
- Wierzbicka, A. (1996). *Semantics: Primes and Universals*. Oxford: Oxford University Press.
- Wilhelm, A. and Saxon, L. (2010). The 'possessed noun suffix' and 'possession' in two Northern Athabaskan languages. Invited talk presented at the Workshop on Nominal Dependents, Banff, May 8.
- Wilkins, D. P. (1989). Mparntwe Arrernte (Aranda): studies in the structure and semantics of grammar. Ph.D. dissertation. Australian National University.
- Wilkins, D. P. (1991). The semantics, pragmatics and diachronic development of 'Associated Motion' in Mparntwe Arrernte. *Buffalo Working Papers in Linguistics*, 207–257.
- Williams, E. (1981a). Argument structure and morphology. *The Linguistic Review* 1: 81–114.
- Williams, E. (1981b). On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12: 245–274.
- Williams-van Klinken, C. (1999). *A Grammar of the Fehan Dialect of Tetun, an Austronesian Language of West Timor*. Canberra: The Australian National University.

- Wilmanns, W. (1896). *Deutsche Grammatik. Zweite Abteilung: Wortbildung*. Strassburg: Trübner.
- Wilson, J. L. (2011). Infixation in Yeri. Paper presented at the Morphology Workshop, University of Buffalo, April.
- Wilson, S. (1999). *Coverbs and Complex Predicates in Wagiman* (Stanford Monographs in Linguistics, 179). Stanford: CSLI Publications.
- Wiltshire, C. and Marantz, A. (2000). Reduplication. In G. Booij, C. Lehmann, and J. Mugdan (eds), *Morphology. An International Handbook on Inflection and Word-Formation*. Berlin: Walter de Gruyter, 557–567.
- Windsor, J. (1994). Children's comprehension and production of derivational affixes. *Journal of Speech and Hearing Research* 37: 408–417.
- Winters, K. E. (2007). Angiggami. Inuit Stories. <<https://web.chass.utoronto.ca/inuitweb/stories/>> [Login and Password: guest].
- Wise, H. (1997). *The Vocabulary of Modern French: Origins, Structure and Function*. London: Routledge.
- Wolf, M. (2007). For an autosegmental theory of mutation. In L. Bateman, M. O'Keefe, E. Reilly, and A. Werle (eds), *University of Massachusetts Occasional Papers in Linguistics 32: Papers in Optimality Theory III*. Amherst: GLSA, 315–404.
- Wolf, M. A. (2008). Optimal interleaving: serial phonology–morphology interaction in a constraint-based model. Ph.D. dissertation. University of Massachusetts, Amherst.
- Wolf, M. A. (forthcoming). Candidate chains, unfaithful spell-out, and outwards-looking phonologically-conditioned allomorphy. *Morphology*.
- Wolfenden, S. N. (1928). The prefix m- with certain substantives in Tibetan. *Language* 4(4): 277–280.
- Wolfenden, S. N. (1929). *Outlines of Tibeto-Burman Linguistic Morphology*. London: Royal Asiatic Society.
- Wolff, J. U. (1973). Verbal inflection in Proto-Austronesian. In A. B. Gonzalez (ed.), *Parangal kay Cecilio Lopez: Essays in Honor of Cecilio Lopez on his Seventy-fifth Birthday* (Linguistic Society of the Philippines Special Monograph Issue no. 4). Quezon City: Linguistic Society of the Philippines, 71–91.
- Wood, E. J. (2007). A semantic typology of pluractionality. Ph.D. dissertation. University of California, Berkeley.
- Wood, E. J. and Garrett, A. (2001). The semantics of Yurok intensive infixation. In J. Castillo (ed.), *Proceedings from the Fourth Workshop on American Indigenous Languages*. Santa Barbara: Department of Linguistics, UC Santa Barbara, 112–126.
- Wood, S. N. (2006). *Generalized Additive Models*. New York: Chapman and Hall/CRC.
- Woodbury, A. (1981). Study of the Chevak dialect of Central Yup'ik Eskimo. Ph.D. thesis. University of California, Berkeley.
- Wright, S. (2001). Internally caused and externally caused change of state verbs. Ph.D. dissertation. Northwestern University, Evanston, IL.
- Wunderlich, D. (1997). Cause and the structure of verbs. *Linguistic Inquiry* 28: 27–68.
- Wurm, L. H. (1997). Auditory processing of prefixed English words is both continuous and decompositional. *Journal of Memory and Language* 37: 438–461.
- Wurm, L. H. and Aycok, J. (2003). Recognition of spoken prefixed words: the role of early conditional root uniqueness points. In R. H. Baayen and R. Schreuder (eds), *Morphological Structure in Language Processing*. Berlin: Mouton de Gruyter, 259–286.
- Wurm, L. H. and Ross, S. E. (2001). Conditional root uniqueness points: psychological validity and perceptual consequences. *Journal of Memory and Language* 45: 39–57.

- Wysocki, K. and Jenkins, J. R. (1987). Deriving word meanings through morphological generalization. *Reading Research Quarterly* 22: 66–81.
- Yallop, C. (1982). *Australian Aboriginal languages*. London: Andre Deutsch.
- Yallop, C. (1993). The structure of Australian Aboriginal languages. In M. Walsh and C. Yallop (eds), *Language and Culture in Aboriginal Australia*. Canberra: Aboriginal Studies Press, 15–32.
- Yamanouchi, K. (1992). On the use of be- verbs in Old English glosses to the Lambeth Psalter. *Studies in Medieval English Language and Literature* 7: 53–77.
- Yang, C. (2005). On productivity. *Yearbook of Language Variation* 5: 333–370.
- Yeh, M. (2011). Nominalization in Saisiyat. In F. Ha Yap, K. Grunow-Harsta, and J. Worna (eds), *Nominalization in Asian languages: Diachronic and Typological Perspectives*. Amsterdam: Benjamins, 561–588.
- Yeung, N., Nystrom, L., Aronson, J., and Cohen, J. (2006). Between-task competition and cognitive control in task switching. *The Journal of Neuroscience* 26(5): 1429–1438.
- Young, R. (2000). *The Navajo Verb System: An Overview*. Albuquerque: University of New Mexico Press.
- Young, R. W. and Morgan, W. Sr. (1987). *The Navajo Language: A Grammar and Colloquial Dictionary*. Rev edn. Albuquerque: University of New Mexico Press.
- Young, R. W., Morgan, W. Sr. with the assistance of S. Midgett (1992). *Analytical Lexicon of Navajo*. Albuquerque: University of New Mexico Press.
- Yu, A. (2000). Stress assignment in Tohono O'odham. *Phonology* 17: 117–135.
- Yu, A. (2003). The morphology and phonology of infixation. Ph.D. dissertation. University of California, Berkeley.
- Yu, A. (2004). Infixing with a vengeance: Pingding Mandarin Infixation. *Journal of East Asian Linguistics* 13: 39–58.
- Yu, A. (2007a). *A Natural History of Infixation*. Oxford: Oxford University Press.
- Yu, A. (2007b). The phonology–morphology interface from the perspective of infixation. In M. Miestamo and B. Wälchli (eds), *New Challenges in Typology: Broadening the Horizons and Redefining the Foundations*. Berlin: Mouton de Gruyter, 35–54.
- Yu, A. (2007c). Understanding near mergers: the case of morphological tone in Cantonese. *Phonology* 24: 187–214.
- Zadoks, A. (2000). Switch Evidence in Old Tibetan: between Switch Reference and Evidentiality. Paper presented at the 9th Seminar of the IATS. Leiden University, The Netherlands. June 24–30.
- Zadoks, A. (2002). The Tibetan connection: switch reference and evidentiality from Old Tibetan to Middle Tibetan. Paper presented at the 8th Himalayan Languages Symposium. Bern University, Switzerland. September 19–22.
- Zajceva, M. I. (1978). *Suffiksāl'noe glagol'noe slovoobrazovanie v vepsskom jazyke* [Suffixal verbal derivation in Veps]. Leningrad: Nauka.
- Zandvoort, R. W. (1961). Review of Hans Marchand *Categories and Types of Present-Day English Word-Formation*. *English Studies* 42: 120–124.
- Zandvoort, R. W. (1972). *A Handbook of English Grammar*. 6th edn. London: Longman.
- Zemskaja, E. A. (1978). O paradigmatičeskich otnošenijach v slovoobrazovanii. In N. J. Švedova et al. (eds), *Russkij jazyk. Voprosy jego istorii i sovremennogosostajaniija. Vinogradovskije čtenija* 1–8. Moskva: Nauka, 63–77.
- Zepeda, O. (1983). *A Papago Grammar*. Tucson: University of Arizona Press.
- Zergollern-Miletić, L. (1997). Morphological adaptation of the suffixes of English nouns borrowed in French. *Studia Romanica et Anglica Zagradiensia* 42: 411–416.

- Zhang, L. (1985). The phonetic structure of ABCB type words in modern Lhasa Tibetan. In B. N. Aziz and M. Kapstein (eds), *Soundings in Tibetan civilization*. Delhi: Manohar, 20–34.
- Zide, N. H. (ed.) (1966). *Studies in comparative Austroasiatic linguistics* (Indo-Iranian monographs, 5). The Hague: Mouton.
- Zide, N. H. and Anderson, G. D. S. (2001). The Proto-Munda verb: some connections with Mon-Khmer. In K. V. Subbarao and P. Bhaskararao (eds), *Yearbook of South-Asian Languages and Linguistics 2001*. Delhi: Sage Publications, 517–540.
- Zigmond, M. L., Booth, C. G., and Munro, P. (1991). *Kawaiisu: A Grammar and Dictionary with Texts*. Berkeley: University of California Press.
- Zimmer, K. E. (1964). *Affixal Negation in English and Other Languages: An Investigation of Restricted Productivity*. New York: supplement to *Word*, Monograph 5.
- Zimmermann, E. and Trommer, J. (2011). Overwriting as optimization. *Natural Language and Linguistic Theory* 29: 561–580.
- Zirkel, L. (2010). Prefix combinations in English: Structural and processing factors. *Morphology* 20(1): 239–266.
- Zöfgen, E. (1989). Homonymie und Polysemie im allgemeinen einsprachigen Wörterbuch. In: F.-J. Hausmann, O. Reichmann, H. E. Wiegand, and L. Zgusta (eds), *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Vol. 1. Berlin: De Gruyter, 779–787.
- Zorc, R. D. (ed.), (1986). *Yolngu-Matha Dictionary*. Batchelor NT: School of Australian Linguistics, Darwin Institute of Technology.
- Zwanenburg, W. (1983). *Productivité morphologique et emprunt* [Linguisticae Investigationes Supplementa 10]. Amsterdam/Philadelphia: Benjamins.
- Zwanenburg, W. (1992). Compounding in French. *Rivista di Linguistica* 4(1): 221–240.
- Zwicky, A. M. (1995). Why English -ly is not inflectional. *Chicago Linguistic Society* 31: 523–535.
- Zwicky, A. and Pullum, G. (1987). Plain morphology and expressive morphology. General Session and Parasession on Grammar and Cognition, *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 330–340.

LANGUAGE INDEX

- Acehnese, reduplication 171
- Achooli, verbal derivation 604
- Afar, adjectives 581
- Afrikaans, reduplication 449
- Afroasiatic languages 573–5
 - adjectival derivation 581–2
 - adverbial derivation 582–3
 - nominalization 576–9
 - root-and-pattern morphology 575–6, 579
 - verbal derivation 580–1
- Aghem, evaluative morphology (EM) 309
- Agta, reduplication 555, 556
- Ahtna 670, 671
 - evaluative morphology (EM) 682
 - kinship terms 673
 - nominalization 672, 674
 - verbal derivation 678
 - verbs 685–7, 696–7, 700
- Ainu, evaluative morphology (EM) 312–13
- Aka-Bo 149
- Akan, evaluative morphology (EM) 309
- Akkadian
 - adjectival derivation 582
 - nominalization 577, 578
- Alabama
 - imperfective form 209
 - infixation 147
 - meaning in infixation 144
 - plural formation 206–7
 - reduplication 174, 176
- Alagwa, adverbs 583
- Algonquian languages, evaluative morphology (EM) 315
- Altaic languages 493–6
 - word formation 496–508
- Amele, reduplication 182, 184
- Amharic, verbal derivation 581
- Ammani-Jordanian Arabic, hypocoristic forms 196–7
- Anywa 598, 601–2
- Ao
 - evaluative morphology (EM) 312
 - reduplication 176
- Arabana, compounding 655
- Arabic
 - adjectival derivation 280, 582
 - broken plurals 204
 - Classical, root-and-pattern morphology 575
 - hypocoristic forms 196–7
 - nominalization 577
 - processing studies 107–8
 - templatic morphology 194
- Aramaic, nominalization 579
- Arara
 - infixation 142–3, 145
 - meaning in infixation 144
- Archi, infixation 147
- Arin 509–12
- Armenian 449
 - category-changing 454
 - category-maintaining affixes 456
 - evaluative morphology (EM) 457
 - plural formation 231–2
- Aroma, reduplication 174
- Arosi, reduplication 173
- Arrernte 655
 - Mparntwe, kinship terms 667
 - verb-splitting 148n
- Asilulu, reduplication 555
- Atayal 149
 - infixation 141–2
- Athabaskan languages 669–70
 - affix-ordering 371, 373
 - category maintaining 679–84
 - evaluative morphology (EM) 681–4
 - nominalization 672–3, 674–6
 - roots and stems 670–3

- Athabaskan languages (*Cont.*)
 verbal derivation 677–9
 verbs 684–700
Atsi, affixation 629
Au, infixation 148
Australian languages, meanings of affixes 135
Austronesian (AN) languages 545–6
 affixation 546–50
 infixation 146
 compounding 551–2
 prosody 551
 reduplication 552–7
 stability over time 149–50
 subtractive morphology 550
Avokaya 597

Bade, nominalization 585
Baffin (Inuit dialect) 713, 714, 715, 717,
 719–20, 722
Bafut, evaluative morphology (EM) 309
Bāgandji, evaluative morphology (EM) 310
Bahasa Indonesia *see* Malay
Bahnar
 aspectuals 531
 negation 533
 nominalization 529
 pronominals 531
 reduplication 540
Balese 599
Baltic languages 444
Banjarese, subtractive morphology 550
Bantu, reduplication 182
Bantu languages 559n, 562, 564
 affix ordering 370
 influence on Nilotic languages 770–1
 reduplication 184, 186–8
Barupu, infixation 148
Basque
 comparatives and superlatives 287
 denominal adjectives 279
 influence from Gascon and French 773
 influence from Romance languages 771
 neoclassical compounds 35
Begak, infixation 143
Beja, nominalization 577, 578
Bella Coola, lexical affixes 39, 40
Bengali
 affixation 460–1
 reduplication 459
Berber languages 573, 575
 compounding 578
 root-and-pattern morphology 576
Beria, nominalization 606
Berta, verbal derivation 604
Bhojpuri (Hindi) 445
 affixation 467
 evaluative morphology (EM) 457
Bidiya
 nominalization 586, 587
 reduplication 589
Bikol
 reduplication 176, 553
 subtractive morphology 550
Birom, infixation 148
Bolaang Mongondow, reduplication 553
Bole, infixation 148
Bontok 149
Botolan Sambal, reduplication 553
Boumaa Fijian, reduplication 183–4
Breton
 numerals 333–4
 paradigms 359, 366
Bru
 desideratives 532
 locatives 533
 nominalization 529
 reciprocal affixes 530
 reduplication 539
 verbal derivation 532
Bugan, compounding 542
Bulgarian, affix-ordering 381
Bundjalung, verbal derivation 660
Bunun, reduplication 554–5
Buryat 499, 502, 504

Cantonese, tonal alternation 216
Carrier
 nominalization 676
 verbs 697–9
Catalan
 affixation, semantic change 352
 category-changing 454–5
 category-maintaining affixes 456
 denominal adjectives 279
 evaluative morphology (EM) 296–7
Cavineña, circumfixes 127–8

- Cebuano 150n
 subtractive morphology 550
 Celtic languages 443–4
 Chaco languages 744
 Chadic languages 573
 affixation 583–7
 reduplication 587–9
 Chagatay 497, 500–1, 504–5
 Chaha
 consonant mutation 210–12
 featural affixes 139n
 Chamorro
 meaning in infixation 144, 145
 reduplication 556
 Chechen, reduplication 179–80
 Cheq Wong, reduplication 537
 Chinese
 Cantonese, tonal alternation 216
 Fuzhou, diminutives 301
 Mandarin/Standard 609–11
 affixation 611, 613, 616–20, 625–6
 infixation 146
 compounding 542
 conversion 620
 influence from Manchu 774
 morphosyntactic features 611–13
 reduplication 614–15
 speech production studies 111
 voicing 622
 Pingding 786
 Choguita Rarámuri (CR) 727–9
 argument changing 732–40
 verbal derivation 729–32, 740–2
 Chorote languages 744, 745
 evaluative morphology (EM) 757
 nominalization 749–50, 753, 755–6, 759
 verbal derivation 762–3
 Chrau
 compounding 543
 involuntary prefixes 533
 reciprocal affixes 530
 statives 530
 Chukchi, reduplication 47, 173, 178
 Ciyao, reduplication 186
 Cree, evaluative morphology (EM) 315
 Cupeño, habilitatives 198–9
 Cushitic languages 573
 adverbs 582–3
 nominalization 577–8, 579
 verbal derivation 580, 581
 Cuzco Quechua, causatives 266–7
 Czech
 comparatives and superlatives 287
 compounding 447
 evaluative morphology (EM) 303
 Dagur 499, 502–3, 508
 Dakkhani, reduplication 459
 Danish, compounding 448
 Dëne Sųtıné 684, 695, 697
 Diegueño, evaluative morphology (EM) 310–11
 Dinka 592, 596, 786
 Diola-Fogny, evaluative morphology (EM) 310
 Ditidaht/Nitinaht, reduplication 177–8
 Diyari
 category maintaining 666
 nominalization 659
 reduplication 170
 Dogon 558–9
 Dutch
 affixation
 circumfixation 452
 polysemy in affixation 348
 semi-affixes/affixoids 122
 complex words 110
 compounding 448
 conversion 161, 163
 denominal adjectives 279
 evaluative morphology (EM) 457
 historical change 406
 nominalization 243–4, 250
 paradigms 362–3, 364–5
 processing studies 106, 107
 socio-geographic variation 411–12
 speech production studies 111, 112
 suppletive allomorphy 221–2, 230n
 synthetic compounds 42
 East Chadic
 nominalization 584
 reduplication 589

- East Dangla
 nominalization 586, 587
 reduplication 587, 588, 589
 verbal derivation 581
- East Tucanoan languages
 influence on Hup 775–6
 influence on Tariana 774, 775
- Eastern Armenian, causative/inchoative
 alternation 265
- Egyptian
 adjectival derivation 582
 compounding 578
 nominalization 577
 verbal derivation 580
- Egyptian languages 573, 575, 576
- Eipo, infixation 148
- Enets 474
- Enga, influence on Yimas 774
- English 444, 446–7
 a/an alternation 233
 adjectival derivation 278, 279
 adjectives, relational vs. qualitative 285
 adverbs 295
 predicative 291–2
 affixation 118
 allomorphy 125–6, 227–8
 historical change 384–6
 meanings in 134–5
 ordering 93–4, 116, 131–2, 372, 375, 377–81
 polysemy in suffixation 342–3, 345
 rival suffixes 70–2, 74–5, 80–3
 semantic change 352
 semi-suffixes 30, 31–2
 suffixation 26–7, 116
 unique affixes 121
 ambiguous novel compounds 32–3
 American, numerals 336
 borrowing
 from Greek 767–8
 loanwords in Pama-Nyungan
 languages 657, 661–3
 loanwords in Russian 121–2
 suffixes 76–80
 bound forms 34, 35
 category-changing 11, 454–5
 by apophony 123–4
 causative/inchoative alternation 270–1
 comparatives 203
 and superlatives 287, 288–9
 conversion 44–5, 120–1, 157, 158, 160, 450–1
 Early Modern English (EmodE),
 affixation 392–5, 397
 evaluative morphology (EM) 302, 303, 456–7
 function words 317, 319–21, 324, 327–8,
 331–2
 headedness 133
 hypocoristic forms 201–2
 influence on Finnish compounding 769
 influence on Hindi 470–2
 influence on Mandarin Chinese
 affixation 616
 irregular forms 68–9
 language acquisition 425–9, 430–3, 434–5,
 437
 Middle English (ME) 30
 affixation 384–5, 390–2, 393, 403–6
 neoclassical compounds 35–6, 119–20
 nominalization 235–40, 244–5, 249–52,
 256, 453
 numerals 332–6
 Old English (OE) 29–30
 affixation 384, 385, 386–90, 395, 398,
 400–3
 paradigms 358, 360
 past tense 17
 phonaesthemes 102
 plural formation 20–1, 192
 processing studies 107
 reduplication 47–8, 171, 180
 scientific, meaning in infixation 144, 145
 speech production studies 111, 112
 spoken by Hungarian-speakers 419–22
 stress alternation 122–3, 216
 synthetic compounds 41, 42
 tmesis 137–8
 transition from lexeme to suffix 28, 29–30
 verbal derivation 257–8, 262, 264
- Erzya dialect of Mordvin 481–2
- Eskimo-Aleut languages 702–3
 lexical affixes 40
 reduplication 710
 word formation 703–10
- Estonian 474, 480
- Even 498–501, 504–7

- Evenki, propriative adjectives 504–8
- Ewe
 affixation, semantic change 351
 reduplication 172
- Figuig, nominalization 579
- Fijian
 compounding 552
 reduplication 175
- Filipino 546
- Finnic languages 473–4
- Finnish 474
 adjectival derivation 280, 477–8
 affixation 476–9
 compounding 769–70
 conversion 479–80
 inhabitant suffix 343
 nominalization 476–7
 numerals 334
 processing studies 107
 relationship with Estonian 480
 relationship with Sámi 481
 word-formation 480
- French 339, 444
 affixation
 borrowed suffixes in English 76–80
 ordering 19
 semantic change 352
 category-changing 454–5
 category-maintaining affixes 456
 comparatives 288
 conversion 160, 450–1
 defective paradigms 21
 evaluative morphology (EM) 457
 function words 317
 influence on Basque 773
 influence on Early Modern English (EmodE) 394–5
 language acquisition 429, 433, 434, 435
 neoclassical compounds 35
 nominalization 241, 255, 454
 numerals 335
 Old French (OF) 385
 influence on Middle English (ME) 391–2, 403–6
 orthographic production studies 113
 paradigms 355
 polysemy 340
 proclitics 23
 reduplication 47
 transition from lexeme to suffix 28
- Fula, diminutives 25
- Fur 600, 601, 604–5
- Gascon, influence on Basque 773
- Georgian, function words 322n, 324–6
- German 444
 adjectival derivation 278
 affixation
 circumfixes 127
 historical change 384, 386
 interfixes 126
 polysemy in 346
 prefixation 36–7
 semi-suffixes 30–1
 restrictions 372
 Bavarian, contact-induced
 grammaticalization 772–3
 bound forms 34
 category-changing 454
 conversion 159, 163, 449, 451
 evaluative morphology (EM) 457
 historical change 406
 language acquisition 435–6, 437
 Middle High German 29
 neoclassical compounds 36
 New High German 29
 nominalization 238, 242, 244, 255, 453–4
 Old High German 29
 orthographic production studies 112–13
 reduplication 47, 48
 subjects 12
 synthetic compounds 41, 42
 transition from lexeme to suffix 29
 umlauts 212–13
- Germanic languages 444
 compounding 27
 influence on Old English (OE) 386–90
 language acquisition 427, 428, 433
 prefixation 36
 transition from lexeme to suffix 29–30
- Gidar
 nominalization 584, 586
 reduplication 589

- Gilyak/Nivkh, consonant mutation 210
- Giziga
adverbial derivation 583
nominalization 577, 584, 586, 587
reduplication 588, 589
- Greek 444, 447
Ancient 338–9, 443
grammars 14
loanwords in English 767–8
neoclassical compounds 35, 119–20, 385
nominalization 240–1, 242, 250, 255, 453
numerals 333–4
- Guaraní, grammars 14
- Guaykuruan languages 743
- Gumuz, verbal derivation 605
- Gwandara, nominalization 585
- Hadza, infixation 148
- Halang, reduplication 540
- Halkomelem, meaning in infixation 144
- Hamar, nominalization 579
- Hän 670
- Hanunóo, prosody 551
- Hare 670
- Hausa 573
affixation 583
nominalization 242, 584, 585
reduplication 177, 588
- Hdi
nominalization 585
reduplication 587, 588
- Hebrew
adjectival derivation 582
Biblical, root-and-pattern morphology 575–6
Israeli, nominalization 578
language acquisition 430, 433, 434, 437, 438
naïve discrimination learning (NDL) 115
nominalization 579
processing studies 107–8
speech production studies 111
templatic morphology 195–6
verbal derivation 580
- Hiaki, reduplication 188
- Hidasta, pitch alternation 216–17
- Hiligaynon, reduplication 174
- Hindi 445, 447
affixation 460–1
prefixation 461–3
suffixation 463–72
conversion 460
evaluative morphology (EM) 457
reduplication 171–2, 459
- Hittite 443
- Hoava
affixation 548
infixation 137, 138, 139, 141, 143
- Hua, infixation 147
- Hungarian 474, 486–9
affixation 475–6
evaluative morphology (EM) 302
influence on Lovâri Romani 770
influence on Slovak and English as second languages 419–22
language acquisition 430
verbal derivation 350, 487–8
- Hup, influence from East Tucanoan languages 775–6
- Hupa 670
nouns 681
suppletive allomorphy 232–3
- Icelandic
compounding 447–8
language acquisition 433–4
- Ijoid 558–9
- Ilokano, affixation 549
- Indo-Aryan (IA) languages 445–6, 447, 458–72
reduplication 449
- Indo-European (IE) languages 443–6
word formation 446–57
- Indo-Iranian languages 444–6
- Indonesian
causative/inchoative alternation 269
reduplication 171, 175
- Ineseño Chumash/Samala, reduplication 187–8
- Inuit 702–3, 710–12
affixation 712–22
modification 719–22
nominalization 712–14
verbal derivation 714–19
- Inupiaq 709

- Inuttut, Labrador *see* Labrador
 Iraqw
 nominalization 577, 578, 579
 verbal derivation 580
 Irish, verbal nouns 24
 Iroquoian languages, bound roots 38
 Italian 444
 affixation 118–19
 semantic change 352
 contact-induced grammaticalization 772–3
 evaluative morphology (EM) 298, 313–14, 350, 457
 language acquisition 431
 nominalization 453
 reduplication 48, 449
 transition from lexeme to suffix 28
 Itbayaten, reduplication 553
 Ithmus Mixe, featural affixes 139n

 Jah-Hut, nominalization 528
 Jaminjung, compounding 656
 Jamul, circumfixes 128
 Japanese
 evaluative morphology (EM) 301, 312
 hypocoristic forms 201
 prefixes 124
 Shizuoka, emphasis 207–9
 Javanese
 elative forms 213
 Old, reduplication 553
 Jeh, meaning in infixation 144, 145
 Joola Banjal
 nominalization 571
 verbal derivation 562

 K'abeena, nominalization 577, 579
 Kachin, affixation 630
 Kadazan, affixation 549
 Kadu languages 591
 Kalaallisut (Inuit dialect) 705–6, 708–9, 711, 713, 722
 Kalanga, lack of affixation 130
 Kalmyk 502, 504–5
 Kambara, circumfixes 128
 Kanembu, vowel alternation 214
 Kangiryuarmiut (Inuit dialect) 712, 716–17, 718

 Kanuri 592, 597
 nominalization 605–6
 numerals 333–4
 verbal derivation 603, 604
 Kapampangan, reduplication 553
 Karao, evaluative morphology (EM) 311
 Karo Batak, reduplication 553
 Katu 527
 aspectuals 531
 compounding 541–2, 543
 involuntary prefixes 533
 nominalization 529
 reduplication 540
 statives 530
 Kaurna 653–4
 Kavalan 149
 Kazakh 497–9, 501–3
 propriative adjectives 504
 Kazan Tatar, affix-ordering 381
 Kelabit, affixation 548
 Kentakbong
 aspectuals 531
 desideratives 532
 Ket 509–12
 affix-ordering 373–4
 nominalization 513–14
 verb structure 514–19
 Kham, evaluative morphology (EM) 312
 Khanty 474
 Kharia, reduplication 178–9
 Khasi 526
 compounding 542
 desideratives 532
 nominalization 528, 529
 reduplication 537, 539
 statives 530
 Khmer 521
 affixation 125, 151–3, 530
 negation 533
 nominalization 528, 529
 presyllables 524
 Khmu 526
 aspectuals 531
 nominalization 528, 529
 reciprocal affixes 530
 statives 530, 533
 verbal derivation 532

- Khoekhoe, evaluative morphology (EM) 309
Kinande, reduplication 171n, 178
Kinyarwanda, reduplication 186–7
Kipsikiis 593–4, 595, 596
Kiribatese, reduplication 174
Klallam, meaning in infixation 144, 145
Koasati, plural formation 206, 207
Koho
 compounding 543
 nominalization 529
 pronominals 530–1
 reduplication 539, 541
 verbal derivation 532
Kokota, reduplication 175
Kolami, reduplication 171
Koman, verbal derivation 605
Komi/Zyryan 474, 484–6
Koryak, reduplication 173
Kott 509–12, 519
Koyukon 670, 672
 kinship terms 673
 nominalization 672, 674
 verbs 690–1, 694
Kri
 infixation 147
 nominalization 529
 verbal derivation 532
Kujamaat Joola, paradigms 356
Kukatja 655
Kurdish 445, 449
Kwakw'ala, affix-ordering 373, 381–2
Kwazá, reduplication 183–4

Labrador (Inuit dialect) 708n, 710, 714, 716, 718, 720–1, 780
Ladin, contact-induced
 grammaticalization 772–3
Latin 443
 adverbs 292
 affixation 343–4
 prefixation 451–2
 semantic change 352
 Classical, transition from lexeme to
 suffix 28–9
 grammars 14, 21–2
 influence on Early Modern English
 (EmodE) 393–4
 influence on Middle English (ME) 391–2
 neoclassical compounds 35, 119, 385
 numerals 335
 paradigms 356
 polysemy 341
 polysemy in suffixation 343–4
Lele
 nominalization 586–7
 reduplication 588
Lendu 593
Lepcha, affixation 629
Lese 601
Leti, infixation 141
Limbu, affixation 630
Luganda, affixation, allomorphy 219–20, 223–4
Luo 592
 influence from Bantu languages 770–1
 nominalization 602–3
Lushootseed, reduplication 176, 185

Ma'di 600
Maasai 600–1, 607–8
Madi 594
Mafa, word-external conditioning 233–4
Magahi 445
 evaluative morphology (EM) 457
Maithili 445
Majang 594
Maka 744
 evaluative morphology (EM) 757
 nominalization 749–53, 755–6, 760
 verbal derivation 761, 763
Makah Melanau 150n
Makasarese, affixation 549
Mal, reduplication 543
Malay/Bahasa Indonesia 545–6
 affixation 548
 compounding 552
 Modern Sri Lanka, influence from Tamil 772
 reduplication 553
 subtractive morphology 550
Malayalam, evaluative morphology (EM) 312
Malayo-Polynesian languages,
 reduplication 174–5, 180
Maltese
 evaluative morphology (EM) 313
 reduplication 176

- Manam, reduplication 554
 Manambu
 infixation 148
 reduplication 176
 Manchu
 influence on Chinese 774
 Written 497, 499–501, 503
 proprietary adjectives 504
 Mandarin Chinese *see* Chinese
 Mande 558–9
 Mang 523, 526
 compounding 543
 reduplication 538, 539, 540
 verbal derivation 532
 Mangarayi, meaning in infixation 145
 Manggarai 149, 552
 Mani, reduplication 173
 Mansi 474
 Marathi, reduplication 459
 Margi, suffixes 124
 Mari 482–4
 Mari languages 474
 Martu Wangka 655
 Masa
 nominalization 584, 586
 reduplication 589
 Mataguyan languages 743–5
 nominalization 748–61
 verbal derivation 761–6
 Mehri, nominalization 579
 Micmac, evaluative morphology (EM) 311
 Migama/Migaama, nominalization
 579, 586
 Mina
 nominalization 586
 reduplication 174, 588
 Mittimatalik (Inuit dialect) 714, 718
 Miya
 nominalization 585, 587
 reduplication 587, 589
 Mlabri 526
 Mnong, reduplication 537,
 539, 541
 Modgel, nominalization 585
 Mofu-Gudur, nominalization 584
 Mohawk, lexical affixes 39
 Mokilese, reduplication 47, 170
 Mokilko
 nominalization 585, 586
 reduplication 589
 Mon-Khmer languages 520–3, 543–4, 786
 affixation 523–34
 compounding 523, 540, 541–3
 reduplication 523, 534–41, 543
 stability over time 150
 Mon, negation 533
 Mongghul 498
 proprietary adjectives 508
 Mongolian 499
 Mongolic languages 494–5 *see also* Altaic
 languages
 Mordvin (Erzya dialect) 481–2
 Mordvin languages 474
 Mosetén, evaluative morphology (EM) 311
 Motu, reduplication 555
 Muna, evaluative morphology (EM) 312
 Munda languages 520–1, 526
 stability over time 150
 Muong
 compounding 542
 reduplication 539
 Mupun
 nominalization 585
 reduplication 587, 588
 Nanai, proprietary adjectives 504–8
 Nancowry 526, 527
 nominalization 528, 529
 Nandi, adpositional derivation 603
 Navajo 670, 671
 causative/inchoative alternation 271
 evaluative morphology (EM) 681–2
 nominalization 675–6
 nouns 681
 verbs 690, 699–700
 Ndebele, reduplication 186
 Nélémwa, evaluative morphology (EM) 312
 Nenets 474, 489–91
 Ngaju Dayak, reduplication 553
 Ngeq, reduplication 543
 Ngiti 594–5, 599
 Ngiyampaa
 evaluative morphology (EM) 667–8
 nominalization 658–9

- Niger-Congo languages 558–9
 nominalization 571–2
 noun class systems 567–70
 verbal derivation 559–67
- Nilo-Saharan languages 591–3
 affixation 593–608
 case marking 607–8
 category changing 601–6
 category-maintaining 599–601
- Nilotic languages, influence from Bantu
 languages 770–1
- Nimboran, blocking 229n
- Nisgha, lexical affixes 39
- Nitinaht/Ditidaht, reduplication 177–8
- Nivacle 744, 747
 evaluative morphology (EM) 757–8
 nominalization 749–56, 760
 verbal derivation 761–4
- Nivkh/Gilyak, consonant mutation 210
- Noni, infixation 148
- Norogachi Rarámuri 728
- Norwegian
 adverbs 292
 deverbal adjectives 278
- Nuxalk, lexical affixes 39
- Nyangumarta 653, 654, 655
 compounding 655
 reduplication 656
 verbal derivation 660–2
- O’Odham *see* Tohono O’Odham
- Occitan
 affixation, semantic change 352
 evaluative morphology (EM) 313
- Omoti languages 573
 adjectival derivation 581–2
 nominalization 577
 verbal derivation 581
- Oriya, affixation 460–1
- Oromo, verbal derivation 581
- Ouargla
 adverbial derivation 583
 verbal derivation 580
- Paamese
 compounding 551
 reduplication 175
- Pacoh 522
 compounding 542
 involuntary prefixes 533
 locatives 533
 negation 533
 number 534
 pretence 533
 pronominals 531
 reduplication 535, 537
 statives 530, 533
- Palaung 526
 nominalization 529
 number 534
 pretence 533
 reciprocal affixes 530
 reduplication 539
- Palawano, reduplication 556
- Pama-Nyungan languages 651–2
 affixation 658–68
 category maintaining 664–8
 compounding 655–6
 evaluative morphology (EM) 667–8
 loanwords 653–5
 nominalization 658–9
 reduplication 656–7
 verbal derivation 659–64
 word formation 652–3
- Pangasinan, reduplication 555
- Päri 595–6
- Pashto 445
- Passamaquoddy, evaluative morphology
 (EM) 299
- Pazeh 150n
 affixation 548
- Permic languages 474, 484–6
- Pero
 nominalization 584–5
 reduplication 589
- Persian 445, 449, 500–1
 affixation 462
 reduplication 48
- Pévé, nominalization 579
- Pingding Mandarin
 infixation 146, 147
 meaning in infixation 144
- Pipil, influence from Spanish 774–5
- PMP, compounding 551–2

- Polish 444
 affixation, allomorphs 224
 agreement 22–3
 denominal adjectives 279
 evaluative morphology (EM) 313, 315
 inflection 10–11, 17, 21
 language acquisition 433, 436
 neoclassical compounds 35
 paradigms 356
 Portuguese 444
 compounding 448
 conversion 450
 deverbal adjectives 278
 evaluative morphology (EM) 457
 Modern Sri Lanka, influence from
 Tamil 772
 Proto-Athabaskan 672
 Proto-Austronesian (PAN) 149–50, 546–7
 Proto-Chamic, infix borrowing 153
 Pumpokol 512
 Punjabi
 affixation 460–1, 467
 reduplication 459
 Puyuma, stability over time 150

 Quechua, affix-ordering 378
 Quileute, infixation 141

 Raluana 149
 Rarámuri 724, 726–9
 argument changing 732–40
 verbal derivation 729–32, 740–2
 Rawang, affixation 630
 Rengao 526, 530
 Rgyalrong 630–1, 649–50, 786
 affixation 634, 637–49
 apophony 635–7
 category changing 642–6
 category maintaining 637–42
 ideophones 646–9
 Japhug, affixation 629
 nominalization 633, 643–5
 reduplication 633, 634–5
 stem formation 632–3
 suppletion 637
 verbal derivation 645–6, 648–9
 voicing 622, 624

 Riang-Liang 526
 River Guarijío 727–8, 730, 732, 735
 Romance languages 444
 compounding 27
 influence on Basque 771
 language acquisition 427, 429
 nominalization 241, 242, 244
 transition from lexeme to suffix 28–9
 Romani, Lovāri, influence from
 Hungarian 770
 Romanian
 compounding 448
 evaluative morphology (EM) 315
 nominalization 255
 reduplication 449
 Rotuman, templatic morphology 197–8, 199
 Roviana, reduplication 177
 Ruc 521
 locatives 533
 nominalization 528
 pronominals 531
 statives 530
 Rundi, verbal derivation 562
 Russian 444, 474n
 adverbs 295
 evaluative morphology (EM) 297
 influence on Arin 510
 influence on Permic languages 484
 language acquisition 431, 433, 436
 reduplication 48, 181
 semi-affixes/affixoids 121–2
 suppletive allomorphy 226

 Sach 521
 Saharan languages, category changing 605–6
 Saisiyat, nominalization 241
 Sakha, proprietive adjectives 504
 Salar, proprietive adjectives 505
 Salishan languages, lexical affixes 40
 Samala/Ineseño Chumash,
 reduplication 187–8
 Sámi 474, 481
 Samoyedic languages 474
 Sanskrit 443, 447
 affixation 461–3
 function words 323–4, 327–32, 337
 numerals 335

- Sara Mbay 595
Sedang 526
Sekani 670
Semai
 infixation 141
 meaning in infixation 144
Semelai 527
Semitic languages 573, 574
 adjectival derivation 582
 language acquisition 427
 nominalization 578
 root-and-pattern morphology 575–6
 templatic morphology 191, 193–7, 199
Sengoi, nominalization 527
Serbo-Croatian
 category-changing 454
 nominalization 453
Shabo 591
Shughni, function words 324–5
Shuswap, meaning in infixation 145
Siar Lak, evaluative morphology (EM) 312
Sidaama, nominalization 577–8, 579
Sidamo
 nominalization 577
 verbal derivation 581
Sinhala, reduplication 459
Sino-Tibetan languages 609
Slavey
 evaluative morphology (EM) 682–4
 verbs 687–8, 692–3
Slavic languages 280, 444
Slovak
 adverbs 293n
 evaluative morphology (EM) 302, 307, 313–14, 315
 function words 319n
 neologisms 779–80
 nominalization 453
 paradigms 355, 357, 359, 363
 spoken by Hungarian-speakers 419–22
Somali
 adjectival derivation 581
 verbal derivation 580, 581
Somrai, nominalization 585
Spanish 444
 adjectival derivation 277, 280
 adverbs 292–3
 affixation
 infixation 452
 polysemy in suffixation 345
 semantic change 351, 352
 causative/inchoative alternation 267–9
 comparatives and superlatives 287
 conversion 160
Dominican
 diminutives 301
 reduplication 181
evaluative morphology (EM) 457
identical adverbs and adjectives 158–9
influence on Pipil 774–5
loanwords in Wichi 758
Mexican
 diminutives 301
 reduplication 181
nominalization 287–8, 453
Peninsular, evaluative morphology (EM) 350
reduplication 48, 449
relational vs. qualitative adjectives 285
sociolinguistics 408
systematic properties of adjectival
 suffixes 281–3
transition from lexeme to suffix 28
Spokane, lexical affixes 39
Sre
 reduplication 539
 statives 530
Supyire, evaluative morphology (EM) 298–9
Sura, nominalization 585
Swahili
 affixation 130
 evaluative morphology (EM) 310
 verbal derivation 562
Swedish
 compounding 448, 449
 denominal adjectives 279
 deverbal adjectives 278
Sye, reduplication 556
Syriac, adjectival derivation 582
Tagalog 150n, 547
 affix-ordering 376–7
 circumfixes 128
 prosody 551
 reduplication 174, 188, 554
Tajik 449

- Tama, verbal derivation 604
 Tamashek *see* Tuareg
 Tamil
 influence on Malay and Portuguese
 creoles 772
 reduplication 47
 Tanacross
 nominalization 673, 675
 nouns 681
 Tapiete, evaluative morphology (EM) 311
 Tariana, influence from East Tucanoan
 languages 774, 775
 Tarok, reduplication 173
 Tashelhiyt
 adjectives 581
 nominalization 577
 Tatar, infixation 146n
 Tausug, infixation 141
 Temiar
 aspectuals 531
 reduplication 537
 Terena, nasal harmony 214–15
 Tetun, meaning in infixation 145
 Thai
 evaluative morphology (EM) 312
 infixation borrowed from Khmer 151–3
 meaning in infixation 144, 145
 Thao
 affixation 549
 reduplication 552–3, 554, 555–6
 triplication 556–7
 voice 547
 Tibetan 620–1
 affixation 623–30
 compounding 623, 625
 nominalization 624–8
 phonology 621–2
 reduplication 622
 Tigak, reduplication 175
 Tigre, nominalization 579
 Tiipay, circumfixes 128
 Till, reduplication 174
 Timugon Murut, subtractive morphology 550
 Tiv, tonal alternation 215
 Tɬɨchɔ Yatii 684
 Toba Batak 149, 551
 Tohono O'Odham
 causatives 266–7
 deletion 192, 205–6, 218
 reduplication 170
 Tok Pisin, reduplication 554
 Tolowa 670
 Tongan 547
 causative/inchoative alternation 265
 Ts'amakko, adjectives 581, 582
 Tswana
 nominalization 571–2
 noun class systems 567–70
 verbal derivation 559–62, 564
 Tuareg/Tamashek
 adjectives lacking 581
 nominalization 578
 plural forms 213–14
 Tungusic languages 495, 505 *see also* Altaic
 languages
 Turkana 597–8
 case marking 608
 nominalization 602
 Turkic languages 494 *see also* Altaic languages
 Turkish 497–503
 affixation 17
 agreement 22–3
 proprietary adjectives 504–5
 reduplication 48
 suppletive allomorphy 225, 226
 Tuscarora, bound roots 38
 Tzutujil, circumfixes 128
 Ubangian 558–9
 Udihe 498–502
 Udihe, proprietary adjectives 504–7
 Udmurt/Votyak 474, 484–6
 Ugric languages 474
 Ukrainian 444
 Ulwa
 prosody 203
 verbal derivation 274
 Unangax/Unangam Tunuu 702, 705–7
 Uncunwee, case marking 607
 Upper Kuskokwim, verbs 685, 699
 Upper Sorbian
 denominal adjectives 280
 evaluative morphology (EM) 313
 Uralic languages 473–5
 word formation 475–91
 Urdu 445

- Utkuhiksalingmiut (Inuit dialect) 712–13,
714–16, 717–20, 722
- Uto-Aztecan (UA) languages 724–7
argument changing 732–40
conversion 731
instrumental prefixes 727–9
verbal derivation 729–32, 740–2
- Uummarmiut 708, 710
- Uyghur 497–500, 502–3
- Uzbek 502
- Vanuatu, reduplication 554
- Vietnamese 521, 523
compounding 542–3
lack of affixation 7
nominalization 527
pronominals 531
reduplication 535–41
Southern, function words 322n
- Votyak *see* Udmurt
- Wagiman, compounding 656
- Walmajarri, category maintaining 665
- Wandala, nominalization 584, 585–6
- Wangkajunga 655
nominalization 659
verbal derivation 660–2
- Warlpiri
category maintaining 665
causative/inchoative alternation 266, 272
compounding 655–6
nominalization 659
reduplication 47
verbal derivation 660–3
- Warumungu 655
category maintaining 665–6, 667
compounding 656
- West Mudburra 655
- Wetan 149
- Wichi 744, 745–6
evaluative morphology (EM) 757–9, 761
nominalization 749–56, 760–1
verbal derivation 761–6
- Witsuwit'en
evaluative morphology (EM) 682–3
nominalization 672, 675, 676
nouns 680, 681
verbal derivation 678
verbs 689, 692, 693–8
- Wolaitta, verbal derivation 580, 581
- Woleaian, reduplication 174
- Wolof, verbal derivation 562–7
- Yagaria, infixation 147
- Yanyuwa, noun class replication 775
- Yaqui 732n
- Yawelmani
prosody 193, 200–1
reduplication 175
- Yeniseian languages 509, 786
word formation 510–19
- Yeri, infixation 148
- Yimas, influence from Enga 774
- Yir Yoront
category maintaining 665–6
compounding 656
verbal derivation 662–3
- Yoruba 786
infixation 146n
lack of suffixation 8
- Yugambeh Bundjalung 655
- Yugh 509–12
- Yupik, Central Siberian 704–5, 710, 718n
- Yup'ik, Central Alaskan 705–6
affix-ordering 373
lexical affixes 39
- Yupik languages 702, 703
- Yurok
infixation 138–9, 140, 143
reduplication 173
- Zenaga, nominalization 579
- Zulu, evaluative morphology (EM) 315
- Zyryan *see* Komi

NAME INDEX

- A., T. T. 100
 Abad, M. 555
 Abbi, A. 458
 Ackema, P. 50
 Adams, A. 579, 587
 Adams, D. Q. 444
 Adams, V. 35, 45, 155, 161n, 167n, 433
 Agnihotri, R. K. 462n, 463, 467
 Ahland, C. 605
 Aikhenvald, A. Y. 3, 148, 176, 599n, 637, 769, 774, 775–6
 Aimard, P. 433, 436
 Ajello, R. 449, 454, 456–7
 Akinlabi, A. 139n, 209, 214
 Alamin, S. 600
 Alber, B. 203n, 218
 Alderete, J. D. 171, 192
 Alexiadou, A. 236, 240, 241, 244, 246–7, 248, 253, 254, 255, 262n, 267
 Alho, I. 476, 478, 480
 Alhoniemi, A. 483
 Alio, K. 586, 587
 Allen, M. 6, 36, 41, 42, 93, 375
 Allott, A. 609
 Alpher, B. 656, 658, 660, 662–3, 665
 Alsina, A. 260, 371, 742
 Álvarez, C. J. 113
 Alves, M. J. 526, 533, 786
 Amador Rodríguez, L. A. 344
 Amberber, M. 656
 Ambrazas, S. 308
 Ameka, F. 172
 Amery, R. 653, 657
 Anagnostopoulou, E. 262n, 267
 Anbessa, T. 577–8, 579
 Anceaux, L. C. 229n
 Andersen, P. K. 621
 Andersen, T. 595–6
 Anderson, G. D. S. 521
 Anderson, S. R. 6, 16, 18, 22–4, 25, 54, 113, 114, 172, 177, 192, 205, 260, 282, 371, 373, 381, 726n
 Andersson, E. 448, 449
 Andersson, M. 98
 Andrews, S. 105
 Andrus, M. 436
 Anon. 151
 Ansaldo, U. 164
 Anshen, F. 73, 76–8, 410
 Anward, J. 155, 157, 158, 162n, 163, 164
 Appah, C. K. I. 308, 309
 Appiah Amfo, N. A. 308
 Applegate, R. 187
 Appleyard, D. 581
 Apresjan, J. D. 349
 Arad, M. 161, 246, 258, 263
 Archangeli, D. 193, 200
 Arenas, P. 745, 759
 Arndt-Lappe, S. 203n, 218
 Arnott, D. W. 25
 Arntz, R. 13
 Aronoff, M. 6, 7, 16, 18, 20, 35, 68, 70, 71, 72, 73, 76–83, 177, 187, 188, 228, 230, 256, 282, 310, 350, 370, 372, 410, 437, 780
 Aronson, J. 115
 Artiagoitia, X. 35
 Asher, N. 253
 Asher, R. E. 312
 Audring, J. 3, 5, 50
 Augst, G. 436
 Augustinova, M. 96, 99
 Austin, P. K. 170, 659, 666
 Avrorin, V. 495
 Axelrod, M. 691, 693–4
 Aycock, J. 106, 107, 109, 116
 Azeb Amha 577, 581, 582

- Baayen, R. H. 6, 32, 73–6, 87, 89, 91, 92, 93, 96, 97, 99, 100, 101, 102, 103, 104, 105, 106, 107, 109, 110, 111–12, 114–15, 116, 117, 225, 341, 375, 377–8, 380, 410, 411–12, 427, 780
- Badecker, W. 92
- Baerman, M. 3
- Baeskow, H. 396
- Bakema, P. 312, 782, 784
- Baker, B. 656
- Baker, M. 156n, 157, 162n, 166, 235, 246, 248, 260, 262n, 274, 290, 371, 728, 742
- Bakker, D. 778
- Bakker, P. 772
- Balakaev, M. 495
- Balling, L. 101, 109
- Balota, D. 96, 99
- Bamgboṣe A. 146n
- Banker, E. M. 529, 531, 533, 540
- Banksira, D. P. 211, 212
- Barker, C. 91, 247, 250
- Barmič, M. J. 489
- Barnes, J. 265, 266, 273
- Barreteau, D. 233, 584
- Bartee, E. 631
- Bartens, R. 485
- Bartsch, A. J. 98
- Barz, I. 30–1, 121, 389
- Baskakov, N. A. 494
- Bassène, A.-C. 562, 571
- Bat-El, O. 191, 195, 196n
- Bates, E. 426, 431
- Bauer, A. 396, 436
- Bauer, C. 150, 529
- Bauer, L. 3, 6, 18, 21, 22, 35, 36, 67, 68n, 69, 70, 76, 87, 90, 118, 126, 128, 155, 158, 160, 162, 165, 167, 175, 289, 297, 300, 315, 358, 360–1, 386, 410, 438, 450, 452, 455, 782, 784, 786
- Bavelier, D. 116
- Baxter, W. H. 624, 629
- Beard, R. 6, 18, 24–5, 50, 102, 111, 156n, 236, 258, 277, 281, 453, 454
- Beauvillain, C. 430, 439
- Beavers, J. 261, 267n, 268n
- Beck, D. 258
- Becker, J. A. 431
- Becker, T. 320
- Beckman, J. 171
- Beckwith, C. 620, 622, 625n
- Belleti, A. 260n
- Bender, M. L. 591
- Bendor-Samuel, J. 214
- Benedict, P. K. 629, 630
- Benjamin, G. 531, 537
- Bentin, S. 108
- Benua, L. 171
- Benveniste, É. 344–5
- Benzing, J. 495
- Bergen, B. K. 102
- Bergenholtz, H. 156n, 161
- Bergs, A. 385n
- Bergsland, K. 703, 705, 706–7
- Berko, J. 420n, 431
- Berman, R. A. 426, 427, 430, 431, 433, 434, 436, 438
- Bermúdez-Otero, R. 191, 192, 200
- Bertram, R. 32, 91, 97, 102, 105, 106, 107, 116, 436
- Bessé, B. de 11, 12, 18, 23
- Beyer, S. 623, 624, 625, 626–8, 629, 630
- Bhat, D. N. 258
- Bi, Y. 111
- Bick, A. 108
- Bickel, B. 742
- Biddle, N. 652
- Bielec, D. 21
- Bielmeier, R. 622, 624
- Bien, H. 111, 112
- Biese, Y. M. 166n
- Bird, H. 114
- Bisang, W. 769
- Bisetto, A. 253, 254
- Bittner, M. 718, 722
- Blackings, M. 594, 600
- Blake, B. J. 652
- Blake, F. R. 128
- Blažek, V. 493
- Blevins, J. 113, 138, 139n, 140, 141, 143, 148, 170, 175, 177, 786
- Bloch, B. 14–15
- Bloch-Trojnar, M. 24
- Bloomfield, L. 5n, 14–15, 41, 68, 101, 118, 611
- Blount, B. G. 436
- Blust, R. 146, 149, 150, 182, 545, 547, 548, 550, 554, 784, 785

- Bobaljik, J. D. 287
 Bochner, H. 784
 Bodrogligeti, A. J. E. 495
 Bogoras, W. 173
 Bogoyavlenskiy, D. N. 431
 Böhl, A. 111
 Böhmerová, A. 308
 Boldyrev, B. 495
 Bolozky, S. 308, 439
 Bölte, J. 111
 Bonin, P. 96, 99
 Booij, G. 7, 17, 18, 42, 43, 50, 63, 90, 121, 173, 175, 192n, 200, 217, 221, 237–8, 239, 242–5, 246, 248, 250–1, 300, 348, 357, 362–3, 375, 488, 742, 750
 Booth, C. G. 729n
 Bopp, F. 335
 Borer, H. 253, 254, 262, 285
 Bornemann, E. 14
 Borowiak, T. 470, 471
 Borsley, R. 254
 Bosák, J. 360–1, 363, 364
 Botha, R. 41–2
 Boudelaa, S. 108
 Bouchier, L. 132
 Bowern, C. 652
 Boyd, J. 72
 Bozic, M. 104
 Brainard, S. 311
 Bram, B. 160n
 Braunstein, J. 744–5
 Bréal, M. 339, 342
 Bredin, H. 54
 Brekle, H. H. 436
 Bresnan, J. 12, 89, 94, 260, 262, 274
 Briggs, J. 705, 709n, 710
 Brinton, L. 385n
 Brömser, B. 779
 Brooks, P. J. 426, 431
 Broschart, J. 258
 Brovetto, C. 101
 Brown, D. 68, 72, 777
 Brown, L. 308
 Brown, S. W. 341–2,
 Bryan, M. A. 601, 604
 Bryant, P. 427, 431, 433
 Brysbaert, M. 96, 99, 117
 Buchanan, L. 33
 Bui Chi 542
 Burani, C. 105, 106
 Burnley, D. 391, 392
 Butterworth, B. 91
 Buzássyová, K. 360–1, 363, 364
 Bybee, J. L. 134, 147, 182, 187, 293, 372, 637, 680, 742
 Bye, P. 190, 191–2, 205, 207n, 210, 213–14, 217
 Caballero, G. 729n, 731, 732, 740–1, 742
 Cairns, H. 72
 Campbell, L. 725, 769, 774
 Caramazza, A. 32, 92, 105, 106
 Carlson, R. 299
 Carochi, H. 741n
 Carol, J. 745, 763
 Carp, J. 98
 Carpenter, K. L. 427, 434
 Carramazza, A. 111
 Carreiras, M. 104, 108
 Carrier, J. 188
 Carstairs, A. 228–9, 710
 Carstairs-McCarthy, A. 50, 52n, 159n, 202, 259
 Casaus, M. 726
 Castrén, M. A. 509
 Caterina, M. 32
 Cense, A. A. 549
 Centineo, G. 267
 Cerin, M. 666
 Cermele, A. 105
 Chang Kun 629
 Chang, B. S. 623
 Chao, Y. R. 610, 613, 617, 620
 Chapman, D. 395, 410, 412–19
 Chappell, H. 774
 Charekov, S. L. 495
 Charles, T. 144n
 Charney, J. O. 728n
 Chebanne, A. 309
 Chen Zhen 634n
 Chierchia, G. 267
 Childs, G. T. 173, 649n
 Chmielewska, M. 436
 Chmura-Klekotowa, M. 433, 436, 437
 Cho, S. 72

- Chomsky, N. 15, 51, 68, 156n, 192, 260, 262, 291, 358, 374, 410, 711
Christianson, K. 104, 108
Chung, K. S. 612, 617, 620, 786
Chung, S. 258, 716
Churchward, C. M. 197
Cincius, V. 495, 505,
Cinque, G. 707
Ciszek, E. 395
Clackson, J. 446
Clapp, A. 104
Clark, E. V. 258–9, 357, 425–6, 427, 428, 429–30, 431, 433, 434, 436, 437, 438, 439
Clark, H. H. 258–9, 428, 438
Clarke, D. D. 338
Claudi, U. 349, 351
Clauson, G. 494
Coblin, W. S. 621
Cohen, J. 115
Cohen, S. R. 427, 439
Cohn, A. 171
Colangelo, A. 33
Cole, P. 430, 439
Coleman, D. 668
Collinder, B. 475
Comfort, J. 607
Compton, R. 711, 715n, 719–20, 721–2
Comrie, B. 165n, 237, 495, 744
Conathan, L. 179–80
Condoravdi, C. 188
Conrady, A. 623, 624, 629, 630
Contini-Morava, E. 310
Cook, C. 705, 709n, 710, 711, 717n
Cooper, D. 150
Cooper, J. S. 540
Cooreman, A. 639
Corbett, G. 156n, 163, 173, 280
Corbin, D. 20, 381
Corris, M. 148
Corston-Oliver, S. 177
Cortese, M. 96, 99
Coseriu, E. 5
Costa, S. R. 436
Costello, N. A. 527, 531, 533
Coupe, A. R. 312, 631
Cowie, C. 385n
Craig, M. 174
Crass, J. 577, 579, 606
Crazzolara, J. P. 604
Creider, C. A. 603
Creider, J. T. 603
Creissels, D. 4, 5
Crick, F. H. C. 114
Crippen, J. 669
Croft, W. 162, 163, 166, 258, 266, 284, 291
Crowhurst, M. 198, 199n
Crowley, T. 551, 556, 660, 668
Cruse, D. A. 339–40
Crystal, D. 162
Csató, Á. 494
Cueni, A. 89
Culicover, P. W. 430, 434, 439
Cúneo, P. 745
Cusic, D. 173
Cusihuaman, A. 266–7
Cutler, A. 91, 129, 430, 438, 439
Cuyckens, H. 395, 396, 397–9
Czaykowski-Higgins, E. 144n

Dabrowska, E. 426
Dahl, Ö. 769
Daltas, P. 408
Dalton-Puffer, C. 28, 156n, 391n, 392, 395, 403, 406
Dambacher, M. 98, 99
Danks, D. 114
Darmesteter, A. 340, 386n
Dasgupta, P. 460, 462
Davidson, D. 248
Davies, M. 90
Davis, K. 137, 548
Davis, M. 91, 100, 104, 107
Davis, S. 192n, 196–7, 207–9, 217, 778
Davliyeva, A. R. 381
Dayley, J. P. 128, 727
de Almeida, R. 32–3
De Belder, M. 710
de Groot, C. 317
de Guzman, V. P. 155, 158
De Jong, N. H. 107, 225
de Reuse, W. 703, 704, 705, 706, 710, 711, 722
De Schutter, G. 448, 452
De Souza, I. C. 142, 144n, 145
de Vaan, L. 91, 341

- de Vooy, C. G. N. 436
 Deacon, S. H. 427, 431, 433
 DeFrancis, J. 610
 DeLancey, S. 631
 Dell, G. 111
 Delmas, C. 395, 403
 Demers, R. 258
 Denwood, P. 623, 625
 Deny, J. 494
 Derkach, T. 308
 Derwing, B. 32–3
 Derzhanski, I. A. 308
 Descoeudres, A. 436
 Detges, U. 28
 Deutsch, A. 107, 108, 111
 Deutsch, W. 427, 434
 Devlin, J. T. 104
 Di Sciullo, A.-M. 86, 295, 358
 Diakonoff, I. M. 574, 576
 Diependaele, K. 96, 99, 104
 Dietiker, S. R. 288
 Dietz, K. 385, 387–9, 395
 Diffloth, G. 147, 151, 520, 528, 529, 532, 535, 537, 538
 Dijkstra, A. 99, 101
 Dijkstra, T. 91, 96, 105, 106, 109, 111, 116, 225
 Dimigen, O. 99
 Dimmendaal, G. J. 558, 591, 592, 594, 596–7, 600, 602, 603, 607, 608, 770, 786
 Dinh L. T. 537, 539, 541
 Dionysios Thrax 24
 Dirven, R. 339
 Dixon, R. M. W. 183, 604n, 639, 652, 664n, 769, 775
 Đoàn Văn Phúc 529, 531, 532, 533
 Doerfer, G. 494, 495
 Dokulil, M. 5–6, 154, 156, 167, 302–3, 356, 359, 363, 364, 413
 Don, J. 155, 156, 157, 161, 163, 164, 167
 Donaldson, B. 449
 Donaldson, T. 659, 667–8
 Donegan, P. J. 521
 Dorais, L.-J. 703
 Doron, E. 191, 195
 Dovetto, M. 105, 106
 Downer, G. B. 625
 Downing, L. 185, 186, 217
 Dowty, D. 259, 260n, 261, 262, 265
 Drescher, B. E. 332, 708n
 Dressler, W. 50, 159n, 165n, 177, 203n, 240–1, 296, 308, 350, 354, 357, 410–11, 426, 782
 Drieghe, D. 96
 Dryer, M. S. 165n, 308
 Duanmu, S. 610, 613
 Dubois, J. 433, 437
 Dudas, K. 213
 Dunabeitia, J. 104, 108
 Dunn, M. 173
 Durie, M. 171
 Duvivier, G. 14
 Dybo, A. 493
 Dye, M. 102
 Edwards, A. 174, 586, 588
 Egbokhare, F. O. 649n
 Egerod, S. 142
 Ehret, C. 591
 Eklund, A. 98
 El'konin, D. B. 436
 Elbers, L. 436
 Elchlepp, H. 104
 Elman, J. L. 114
 Embick, D. 231n, 232, 260, 262, 263–4
 Emeneau, M. 171
 Emmerick, R. E. 624n
 Enfield, N. J. 147, 258, 529, 532
 Engelberg, M. 477
 Engels, J. 338
 Epps, P. 775–6
 Erben, J. 29, 31, 386
 Ernestus, M. 109, 110, 111, 112, 114, 410, 412
 Ernst, T. 707n
 Escalante, F. 732n
 Estabrooke, I. V. 101
 Ettinger, S. 308
 Evans, N. 258, 651, 658, 667
 Evsevēv, M. E. 481
 Fabb, N. 6, 93, 131, 239, 372, 375, 376, 378, 594, 600
 Fabre, A. 743
 Fábregas, A. 18, 284
 Fabricus, A. H. 656
 Fanego, T. 156n

- Farrell, P. 136, 156n
Faust, N. 710
Fedorova, O. 426
Feintuch, S. 102, 107
Feldman, L. 96, 102, 104, 105, 107, 108
Fellbaum, C. 163
Ferguson, C. A. 426
Ferrand, L. 96, 99
Figueira, R. A. 436
Filbeck, D. 543
Filipovic Durdjevic, D. 100, 105, 107, 114–15
Fillmore, C. J. 397
Finley, S. 209, 210, 214
Fisher, R. A. 166
Fitch, W. T. 80
Fitzgerald, C. M. 170
Fleischer, W. 30–1, 121, 160, 389
Fleischman, S. 386n
Fleming, C. D. 670
Florez Leyva, M. 188
Fodor, J. 265
Ford, M. 89, 94, 107
Forster, K. I. 91, 96, 99, 103–4, 108
Fortescue, M. 703, 704, 705–9, 711, 713, 714, 717, 718n, 719, 721
Fowler, C. S. 725
Fradin, B. 27
Frajzyngier, Z. 174, 574, 582, 584, 585, 586–90
Francez, I. 261
Francis, B. 101
Francis, G. 100
Frank, T. 98, 104
Franko, Š. 410, 412–19
Fraser-Knowles, J. 668
Frauenfelder, U. H. 32, 106
Fromkin, V. 130
Frost, R. 107, 108
Fudeman, K. 7, 18, 20, 310
Fuhrhop, N. 372
Fuller, E. 153
Furdík, J. 363–4

Gaeta, L. 378, 380
Gage, W. W. 520
Gagné, C. 32
Gahl, S. 112, 113
Galand, L. 575, 576, 578

Galkin, I. S. 483
Gallego, Á. 285
Gamillscheg, E. 343
Ganiev, F. A. 146n
Gardner, A.-C. 396
Garrett, A. 138, 140, 143, 148, 173
Geeraerts, D. 312, 339–40, 782, 784
Geilich, B. 626
Genetti, C. 631
Georg, S. 493, 494, 509, 510n
Gerdtts, D. B. 144n
Gerzenstein, A. 745, 747, 750–2, 755–6, 762–3
Ghomeshi, J. 48, 172
Giakoumaki, E. 408
Giegerich, H. 6, 131, 372, 375
Gil, D. 165n, 322n
Gilboa, O. 108
Gilligan, G. 7, 129, 430, 439, 783, 786
Gillis, S. 426
Gimba, A. M. 148
Giraud, H. 105
Gitcho, N. 102, 115
Givón, T. 157, 162n, 163
Gleitman, H. 422
Gleitman, L. R. 422
Gnanadesikan, A. 171
Goelman, G. 108
Göksel, A. 495
Goldberg, A. E. 397
Goldenberg, G. 580
Goldsmith, J. 193
Golla, V. 232, 670, 678, 681
Golluscio, L. 744
Gong Hwang-Cherng 628
Gonnerman, L. M. 104, 114
González Torres, E. 396
González, H. A. 311
González, L. 744
Gonzalez-Diaz, V. 72
Gooch, A. 308
Good, J. 179–80, 742
Goudswaard, N. E. 143
Gràcia, L. 296–7, 308
Gradin, D. 144n
Gragg, G. B. 576, 578, 581, 582
Grainger, J. 98, 99, 101, 104, 105, 107
Grammatika 481

- Grandi, N. 6, 303–4, 314
 Gray, R. D. 149
 Graziano-King, J. 72, 289
 Greenberg, J. H. 7n, 19, 146, 147n, 371, 493,
 558, 574, 580, 591, 743, 780–1, 783, 784
 Greenhill, S. J. 149
 Gregerson, K. 526, 530
 Grégoire, A. 436
 Gregová, R. 308, 314n, 316
 Grevisse, M. 21
 Grimes, S. 209, 386
 Grimshaw, J. 236, 253, 255, 259, 261
 Grönke, U. 308
 Grossmann, M. 386n
 Grünthal, R. 480
 Gualdieri, B. 754
 Guasch, A. 14
 Guerassimovitch, H. 113
 Guevara, E. 50
 Guilbert, L. 366–7, 368, 437
 Guilfoyle, E. 244
 Guillaume, A. 127
 Gulikers, L. 103
 Gurdon, P. R. T. 530
 Gurubasave Gowda, K. S. 176
 Gustafson, J. 102
 Gustafsson, A. 436
 Gvozdev, A. N. 433, 436
 Györke, J. 474
 Gyurmé, K. 623

 Haas, M. R. 178
 Haase, M. 769, 773
 Haberland, H. 448
 Hagiwara, H. 92
 Hahn, M. 623
 Haig, G. 225
 Häikiö, T. 107
 Haiman, J. 180, 181
 Hajdú, P. 475
 Hakulinen, A. 476, 478, 480
 Hakulinen, L. 476
 Hale, K. 258, 259, 261, 266–7, 270–2,
 290, 659
 Hall, C. J. 129
 Hall, R. 663, 779
 Hallap, V. 481

 Halle, M. 6, 16, 53, 102, 111, 192, 232, 263,
 374, 711
 Haller, F. 621
 Haller, S. 98
 Haman, E. 426, 436
 Handel, Z. 609, 624
 Hanique, I. 112
 Hardy, H. 174, 176, 206–7, 209
 Hargus, S. 672, 675–6, 678, 680, 681, 682–3,
 689, 692–8
 Harley, H. 16, 55, 188, 258, 261, 263, 265, 711n
 Harm, M. W. 114
 Harper, K. 705
 Harris, A. 324–6
 Harris, A. C. 740–1, 769, 774
 Harris, C. 98
 Harris, Z. 5, 84, 85
 Harrison, S. P. 170
 Hartshorne, J. K. 101
 Harvey, M. 656
 Haspelmath, M. 7, 50, 156n, 162, 165n, 258,
 265–6, 269, 272–3, 274, 338, 726, 752, 768
 Hathout, N. 81
 Haugen, J. 151, 188, 725, 728, 729, 732
 Hauk, O. 108
 Haviland, J. 88
 Hawkins, J. 7, 129, 430, 439, 783, 786
 Hay, J. B. 6, 73, 76, 92, 93, 112, 115, 116, 131, 286,
 377, 378–80
 Hay, S. E. 427
 Hayashi, M. 722
 Hayes, B. 555
 Hayvaert, L. 156n
 Healey, P. M. 556
 Heath, J. 213, 733n, 735
 Hecht, B. F. 425, 426, 433, 434, 437
 Heeschen, V. 148
 Heine, B. 319, 349, 351, 768–9, 771, 772–3, 775
 Heinonen, T. R. 476, 478, 480
 Hellwig, B. 584
 Henderson, J. 148n
 Hendrix, P. 100, 101, 105, 110, 114–15, 117
 Hengeveld, K. 258
 Henzen, W. 29, 36, 386
 Hercus, L. A. 310, 655, 656, 657
 Hetzron, R. 574, 576, 581
 Hewson, J. 311

- Heyvaert, L. 245, 251
Hildebrandt, K. 631
Hilkenbach, L. 378, 380
Hill, J. 198, 205–6
Hill, N. W. 621, 622, 627, 630, 725
Hiltunen, R. 388
Hippisley, A. 68, 72
Hoàng, Tuệ 529, 531, 532, 543
Hoàng, Văn Hành 539, 542
Hoàng, Văn Ma 530, 532, 539
Hoberman, R. D. 576, 578, 581, 582
Hockett, C. 5, 19, 52, 155, 156n, 158, 159n, 191
Hoffmann, B. 14
Hoffmann, C. 124, 584
Hofmann, D. 308
Hohenhaus, P. 48, 180
Holcomb, P. J. 98, 104
Holden, J. 695, 697, 699
Holes, C. 574, 582
Holton, G. 673, 675, 681
Homič, L. V. 489
Hopkins, E. 102, 107
Hopper, P. 319, 349
Horecký, J. 303, 360–1, 363, 364
Horgan, D. M. 388
Horoi, R. 173
Horvath, J. 267n
Horwood, G. 205, 206
Howard, P. 687–8
Hualde, J. I. 35
Huddleston, R. 155
Hudson, J. 665n
Huehnergard, J. 574–5
Huffman, F. E. 144n, 145, 151–3, 533
Hummel, M. 159
Hünemeyer, F. 349, 351
Hupp, J. M. 430, 434, 439
Hurch, B. 6, 182
Hutchison, J. 597, 603, 604, 606
Hutchison, K. I. 96, 99
Hyman, L. M. 178, 182, 184, 186, 188, 309, 370, 371, 560
Hyönä, J. 106
Jacobini, C. 159n
Ibanez Moreno, A. 396
Ikegami, J. 495
Inkelas, S. 4, 6, 61, 169–70, 171, 177, 179, 182, 187, 200, 222, 229n, 556, 742
Ioannidis, J. P. A. 100
Isanbaev, N. I. 483
Ito, T. 92
Jaberg, K. 350–1
Jackendoff, R. 11–12, 16–17, 48, 54, 172, 259, 261, 285, 291, 401
Jackson, H. 159n
Jacob, J. 125
Jacobs, A. M. 99, 107
Jacobson, S. 703, 704, 705, 709, 713, 718n
Jacques, G. 609, 621, 622, 624, 625, 629, 634n, 635, 636, 638n, 640n, 643, 646
Jaggar, P. 583, 584, 588
Jakobi, A. 606, 607
Jakubowski, P. 396
Jamison, H. L. 104
Janda, L. A. 281, 283, 447
Janhunnen, J. 495, 499, 501, 503, 504
Janssen, N. 111
Jarceva, V. 495
Jarmulowicz, L. 427
Järvikivi, J. 116
Jeddore, R. 717n, 720–1
Jelinek, E. 258, 732n
Jendraschek, G. 771
Jenkins, J. R. 431
Jenner, P. 525, 529, 530
Jenny, M. 533
Jescheniak, J. D. 98
Jespersen, O. 87
Jetté, J. 672, 673, 674–5, 690–1, 700
Ježek, E. 253
Ji, H. 32
Jin, P. 638n
Joannis, M. F. 114
Johannesson, M. 98
Johanson, L. 494, 767
Johns, A. 4, 705, 708n, 709n, 710, 711, 713, 714–15, 716, 721, 780
Johnson, K. 260
Johnson, M. 180
Johnson, R. 108
Johnston, E. 174, 586, 588
Jones, B. 659, 660, 661–2
Jones, E. 672, 673, 674–5, 690–1, 700

- Jong, F. de 495
 Josephson, C. 98
 Jungraithmayr, H. 579, 585, 586, 587, 589
 Jurafsky, D. 6, 181, 300–2, 305, 309, 312, 349

 Kachru, Y. 459–61, 463, 467
 Kačmárová, A. 308
 Kamp, H. 240
 Kandel, S. 113
 Kangasmaa-Minn, E. 474
 Kaper, W. 436
 Kaplan, L. 703, 709, 713, 718n
 Kari, J. 671, 672, 673–4, 678, 682, 685–7, 689, 696–7, 699–700
 Karnchana, N. 524, 528
 Karpf, A. 296
 Kasik, R. 480
 Kastovsky, D. 6, 50, 105, 121, 124, 155, 159, 161, 385, 387, 388, 389–90, 395, 396
 Kaufman, T. 142n, 151–2, 258, 743, 768, 776
 Kaunisto, M. 395
 Kawamura, M. 92
 Kay, P. 397
 Kaye, A. S. 575
 Kazanina, N. 104
 Kečkéšová, V. 420n
 Keegan, J. M. 595
 Kemmer, S. 162n
 Kempe, V. 426, 431
 Kemps, R. 110, 114
 Kennedy, C. 286
 Kenstowicz, M. 210–12
 Kerleroux, F. 167n
 Kernan, K. T. 436
 Kerslake, C. 495
 Kertz, L. 277n
 Kessler, B. 96, 99
 Keszler, B. 486
 Keuleers, E. 96, 99
 Keune, K. 110, 410, 411–12
 Key, H. 176, 180
 Keyser, S. J. 239, 258, 259, 261, 266–7, 270–2, 290
 Khumalo, D. 315
 Kibrik, A. 680, 685, 691, 695n, 697, 699
 Kiefer, F. 475, 486
 Kimball, G. D. 206
 Kimenyi, A. 186–7
 Kimura, D. 101
 King, K. A. 426, 431
 Kinkade, M. D. 144n
 Kinoshita, S. 96–7, 104, 108
 Kiparsky, P. 6, 36, 93, 155, 178, 188, 222, 223, 227, 229–30, 258, 259, 262, 374, 375, 779
 Kiss, É. 486
 Kiyomi, S. 174–5, 176, 180
 Klamer, M. 128
 Kleiber, G. 346
 Klein, D. E. 341
 Klein, J. 102
 Kleinschmidt, S. 707
 Klembárová, E. 421
 Kliegl, R. 98, 99, 105, 115
 Klingvall, E. 278
 Kneisl, M. 475
 Knuth, D. E. 103, 104
 Knutsson, H. 98
 Koch, H. 652, 656, 666
 Koch, K. 258
 Koch, P. 413
 Koester, D. 108, 111, 112
 Kogan, L. E. 576
 Kohler, K. J. 110
 Koivisto, V. 476, 478, 480
 Komlosi, S. 104
 Koontz-Garboden, A. 258, 259n, 263, 265, 267–8, 269n, 273
 Koops, R. 585
 Koreman, J. 109, 112
 Korhonen, R. 476, 478, 480
 Kornexl, L. 396
 Kornfeld, L. M. 27
 Kornflit, J. 254, 495, 499, 502
 Körtvélyessy, L. 4, 6, 7–8, 140–1, 146n, 237, 242, 304–8, 309, 315–16, 322n, 327n, 373–4, 381, 410, 419–22, 448, 449, 452, 453, 463, 757, 782–4, 786
 Kossmann, M. 575, 577, 579, 580, 581, 583
 Kostic, A. 107
 Kostov, K. 770
 Koul, O. N. 461, 462n, 463
 Koziol, H. 386
 Kratschmer, A. 338
 Kratzer, A. 248, 278, 400
 Krause, S. 178
 Krauss, M. E. 673

- Krejnovich, E. A. 519
Kroeger, P. 174
Krott, A. 116, 225, 363
Kruspe, N. 144n, 527, 537
Kuczaj, S. A. II. 430, 434, 439
Kudła, M. 313, 315
Kugler, T. 108
Kuipers, A. H. 144n
Kumari, T. C. 312
Kuperman, V. 32, 96, 97, 101, 102, 105, 106, 107, 115, 116
Kuprijanova, Z. N. 489
Kuriaki, K. 446
Kurusu, K. 190, 192, 205
Kuryłowicz, J. 158
Kuteva, T. 768–9, 772–3, 775
Kutsch Lojenga, C. 594, 599

Laakso, J. 474, 475, 478, 480, 490
Laalo, K. 308, 426
Labelle, M. 267, 431
Labov, W. 11
Laca, B. 344
Lacey, P. 96, 99
Ladányi, M. 410–11, 489
Ladusaw, W. 716
Lafone Quevedo, S. A. 744
Lahaie, K. 574, 577
Lahiri, A. 430, 439
Laine, M. 436
Lakoff, G. 180, 259, 265, 301, 340
Lalou, M. 625
Lambon, R. 114
Lampitelli, N. 710
Lang, M. F. 457
Langacker, R. W. 249, 251, 725, 726
Langdon, M. 310–11
LaPolla, R. J. 259, 261, 609, 621, 628, 630, 636
Lappe, S. 61, 202–3n
Larson, R. K. 19, 294
Laudanna, A. 32, 105, 106
Laufer, B. 629, 630
Laughren, M. 659
Launey, M. 741n
Lavrent'ev, J. I. 483
Lavric, A. 104
Lazar, O. 474
Lazzeroni, R. 447

Le Bleis, Y. 233
Lee, J. 657
Leer, J. 670–1, 677, 681
Lees, R. B. 5, 15
Lehmann, C. 7, 18, 172
Lehnert, M. 87, 93
Lehr, R. 293
Lehrer, A. 338
Lehtisalo, T. 474
Leisi, E. 163
Leitner, G. 157
Lemhoefer, K. 99, 101, 108
Leopold, W. F. 430
Lepämaa, A.-L. 476
Leslau, W. 581
LeSourd, P. S. 299, 308
Levelt, W. J. M. 98, 111, 112, 225
Levi, J. N. 15
Levin, B. 55, 90, 160n, 239, 240, 244–5, 248, 253–4, 259, 261, 262, 266, 267, 277, 286, 401
Levinson, L. 261, 263
Levy, R. 109
Lewandowska-Tomaszczyk, B. 340
Lewis, D. J. 427, 431, 433, 437
Lewis, G. 92, 104, 225
Lewis, M. P. 443, 446, 494–5, 545, 670, 702, 726
Li, J. 542
Libben, G. 32–3, 96, 105, 106, 116
Lichtenberk, F. 554
Lieber, R. 6, 7, 16, 42, 50, 53, 54, 55–6, 63, 65, 68n, 72, 87, 90, 105, 107, 113, 115, 134, 155, 156, 160, 192, 193, 200, 209, 210, 221, 227, 236, 237–8, 239, 240n, 242–6, 248, 250–1, 252n, 257–8, 259–60, 261, 277, 281, 289, 338, 342, 347, 381–2, 386, 396, 397, 399, 401–3, 410, 424, 427, 438, 449, 452, 455, 750
Lieto, A. 476
Lin, X. 638n, 640n, 644,
Lin, Y. H. 144n, 146,
Lin, Y.-J. 631, 635, 636
Lindsay, M. 78–9, 80–3
Lionnet, A. 728
Lipiński, E. 574, 575
Lipka, L. 155, 156n, 157, 162, 163n
Llinás, R. 342
Lloyd, C. 396
Lo Duca, M. G. 434, 436
Lo, S. 105

- Loftis, B. 96, 99
 Loprieno, A. 575, 576, 582
 Loukotka, Č. 743
 Louw, J. A. 308
 Lowe, R. 708, 710, 712, 717, 718
 Lüdtke, J. 386n
 Lukas, J. 606
 Luo, E. 635
 Lupker, S. J. 108
 Luschützky, H. C. 349, 352, 750
 Lustig, A. 629
 Lüttmann, H. 111
 Lutz, A. 395, 406
 Lý Toàn Thắng 529, 531, 532, 543
 Lydall, J. 579
 Lynch, J. 173
- Maas, U. 171, 178–9
 McCarthy, J. J. 6, 61, 115, 137, 139n, 141, 170, 171, 185, 191, 193, 194–5, 196n, 197–9, 200, 203, 204, 231
 McCawley, J. D. 259, 265
 McClelland, J. L. 103, 114
 McConchie, R. W. 395
 Macdonald, R. R. 548
 Macdonell, A. A. 324
 Macfarland, T. 266
 McGinnis, M. 274, 275
 McGregor, W. 655, 664
 McKoon, G. 266
 McNally, L. 286
 McPherson, L. 223
 McQueen, J. 91, 109, 110, 111, 379
 MacWhinney, B. 430, 436
 Magnusson, R. 156n
 Maiden, M. 119
 Majtinskaja, K. E. 475
 Maki, P. 101
 Malchukov, A. 495
 Malik, A. N. 463
 Mallory, J. P. 444
 Manaster Ramer, A. 493
 Manley, T. M. 530, 539
 Manning, C. 718
 Manova, S. 159n, 370, 381
 Marantz, A. 6, 16, 47, 53, 92, 102, 104, 107, 111, 156n, 170, 232, 246–7, 252, 256, 260, 263, 289, 711
- Marchand, H. 5, 30–1, 34, 41, 44, 87, 121, 155, 157, 158, 160, 161, 171, 258, 385–6, 392, 403, 779
 Marelli, M. 100, 105, 114–15
 Markianova, L. F. 475
 Marsico, E. 611
 Marslen-Wilson, W. D. 104, 107, 108, 109, 430, 439
 Martin Arista, J. 396
 Martin, F. 255
 Martin, J. B. 144n, 206
 Martin, S. E. 124, 495
 Masica, C. P. 457, 458, 459, 470
 Masini, F. 3, 5, 50
 Maticsák, S. 481
 Matisoff, J. A. 301, 308, 309, 312, 524, 537, 541, 630, 631
 Mattes, V. 176, 182, 308
 Matthews, P. H. 11, 16, 21, 23, 113, 740
 Matthews, P. M. 104
 Matthewson, L. 258, 274
 Mayer, F. 13
 Mayerthaler, W. 381
 Mchombo, S. A. 260
 Means, N. 527
 Means, P. B. 527
 Mediano, Z. D. 436
 Megerdooimian, K. 265
 Mei, T.-L. 624, 625
 Meillet, A. 621
 Meir, A. 111
 Mel'čuk, I. A. 680
 Melis, A. 584, 586
 Melissaropoulou, D. 308
 Melloni, C. 253, 254, 349
 Melzi, G. 426, 431
 Mendikoetzea, A. 268
 Menges, K. 495
 Méot, A. 96, 99
 Méresse-Polaert, J. 436
 Merlan, F. 135, 144n
 Merlini Barbaresi, L. 203n, 308
 Messineo, C. 745, 754, 760
 Meunier, F. 109
 Meyer, A. S. 111, 112, 225
 Miceli, G. 32
 Michalove, P. A. 493
 Midgette, S. 690

- Milin, P. 100, 101, 105, 107, 114–15, 117
Miller, A. 128
Miller, D. G. 395
Miller, G. A. 163, 767
Miller, K. 148
Miller, R. A. 495
Miller, W. R. 725, 727–8, 730, 732, 735, 738n
Milne, L. 526, 529, 530, 533, 534, 539
Minkova, D. 221, 395
Mintz, M. W. 553
Mironova, N. 426
Mishra, H. 463
Mithun, M. 38–40, 373, 430, 715, 725, 728, 732
Miwa, K. 96, 105, 106, 116
Mohan, S. 469
Mohanani, K. P. 6, 55, 401
Mohanani, T. 55, 374, 375, 401, 470
Möhlig, R. 395
Mondorf, B. 72, 289
Montes Giraldo, J. J. 436
Montler, T. 174, 176, 206–7, 209
Moravcsik, E. 6, 7, 140, 147, 148, 150, 170, 176, 180
Mordecai, K. 101
Morgan, W. Sr. 675, 681, 690, 699–700
Morris, J. 98, 104
Moscato del Prado Martín, F. 96, 102, 104, 105, 107, 114
Moshi, L. 274
Mourelatos, A. 285
Mous, M. 577–8, 579, 580, 581, 582–3
Mudrak, O. 493
Mugdan, J. 18, 156n, 161
Mühleisen, S. 351
Mulford, R. C. 426, 433, 434, 436
Müller, M. 576, 582
Müller, P. O. 18
Müller-Bardey, T. 274
Munro, P. 144n, 729n
Munson, B. 112
Murphy, G. L. 341, 342
Murre, J. M. J. 114
Mutaka, N. 178
Muysken, P. 151, 378

Nagano, A. 133, 396, 779
Nagano, S. 628
Nagaraja, K. S. 526, 529, 532, 537, 542
Nagy, W. 431, 433
Najlis, E. 743, 745
Narrog, H. 319
Nash, D. 656, 659, 660, 664
Nau, N. 769
Nault, K. 33
Naumann, B. 317, 386
Nedjalkov, I. 495
Nedjalkov, V. P. 265
Neef, M. 41, 43, 158, 161n, 167
Neeleman, A. 50
Neely, J. H. 96, 99
Nekushoeva, S. 325
Nelson, D. L. 96, 99
Nercesian, V. 745, 747, 748, 750–4, 756–61, 763–5
Nerlich, B. 338, 339
Nevalainen, T. 385, 394
Nevins, A. 171, 214
New, B. 91, 96, 99, 100, 104, 117
Newman, A. J. 116
Newman, P. 177, 584, 585
Newman, S. 175
Newmeyer, F. J. 15, 84
Newport, E. L. 116
Ngunga, A. 186
Nguyễn Đình Hoà 531, 535, 537, 539, 542, 543
Nguyễn Hữu Hoàng 526, 529, 530, 532, 538, 539, 540, 542, 543
Nguyễn Tài Cần 537
Nguyễn Văn Khang 542
Nguyễn Văn Lợi 526, 528, 530, 531, 532, 533, 538, 539, 540, 542, 543
Nguyễn Văn Tài 529, 531, 532, 533
Nichols, J. 265, 266, 273, 607, 742
Niebuhr, O. 110
Niemi, J. 116
Niepokuj, M. K. 174, 182
Nieuwenhuis, P. 308, 315
Nikitina, T. 89
Nikolaeva, I. 495, 507, 784
Nishida, T. 624
Nolasquez, R. 198
Noonan, M. P. 631
Norman, J. 774
Norris, D. 96–7, 104, 108, 110, 114
Nottbusch, G. 112–13
Nouguier Voisin, S. 562

- Novikova, K. 495
 Noyer, R. 16, 55, 711n
 Ntelitheos, D. 241, 246, 248, 254
 Nuttall, P. A. 335–7
 Nystrom, L. 115

 O'Connor, P. A. 96, 102, 104, 105
 O'Grady, G. N. 653, 655, 656, 658
 O'Grady, W. 155, 158
 O'Reilly, R. C. 114
 Oakes, M. 668
 Ogura, M. 395
 Ohnheiser, I. 18
 Okada, A. 395
 Okell, J. 609
 Olsen, S. 4, 18, 37, 155, 156, 156n, 160, 163, 167n, 385, 756
 Oltra-Massuet, I. 278
 Omar, A. H. 531, 532
 Orgun, O. 222
 Ortiz de Urbina, J. 35
 Orzechowska, A. 21
 Osada, T. 258
 Osing, J. 575
 Otanes, F. 174
 Ott, D. 308
 Ozawa, K. 101

 Packard, J. L. 611
 Pacor, P. 750
 Pagliuca, W. 182
 Pallier, C. 96, 99
 Palmer, C. C. 175, 396
 Palsgrave, J. 404–5
 Panagl, O. 436
 Pancheva, R. 101
 Panócová, R. 308
 Park, I. 216
 Parker, J. 378
 Parkinson, S. 457
 Parsons, T. 278
 Pashler, H. 98
 Paster, M. 223, 227, 228, 231
 Pastizzo, M. J. 102, 107, 108
 Patrie, J. 493
 Patterson, K. 114
 Paul, H. 36, 349, 386
 Paul, P. 158

 Pauly, É. 340
 Payne, T. E. 88–9, 155, 156n, 157, 159n, 162n, 163
 Pearson, J. 17
 Pelletier, F. J. 286
 Pelligrino, C. C. 611
 Pengitov, N. T. 483
 Pennanen, E. V. 155, 156, 157, 158, 160, 161n, 166n
 Pentland, D. H. 315
 Perea, M. 108
 Pereltsvaig, A. 777
 Perkins, R. 182
 Pesetsky, D. 261
 Peters, H. 395
 Peterson, D. 265, 266, 273
 Peterson, J. 171, 178–9, 274
 Petré, P. 395, 396, 397–9
 Pfau, R. 164
 Phaf, R. H. 114
 Pham, H. 105
 Phan Xuân Thành 529, 531, 532, 533
 Picht, H. 13
 Piepenbrock, R. 103
 Pilotti, M. 96, 99
 Pinker, S. 68, 225, 259, 261
 Pinnow, H. J. 521, 542
 Piñón, C. 272
 Pittman, C. 711, 718, 719
 Plag, I. 6, 28, 42, 46, 61, 68n, 87, 90, 92, 93, 116, 155, 163, 257–8, 289, 342–3, 348, 372, 375–6, 377–8, 379–80, 382, 385, 386, 410, 438, 455
 Plaisier, H. 629n
 Plank, F. 146, 150n, 161, 346–7, 354, 454, 726, 781–2
 Platt, J. T. 668
 Plaut, D. C. 114
 Pluymaekers, M. 112
 Pollard, C. 12, 260, 262
 Pollatsek, A. 106
 Poole, K. 258
 Poppe, N. 493
 Poser, W. 170, 201, 676, 697–9
 Postal, P. 15, 324
 Pou, S. 525, 529, 530
 Pounder, A. 355–6, 357–8, 366–8
 Premsrirat, S. 526, 529, 530
 Prentice, D. J. 550
 Priestley, J. 160n

- Prieto, V. M. 305, 308, 407–8
Prince, A. 141, 170, 171, 185, 193, 198–9, 200, 203, 204, 231
Pulleyblank, D. 215
Pulleyblank, E. G. 622
Pullum, G. 296, 321n
Pulvermüller, F. 108
Pustejovsky, J. 253, 254, 340
Puura, U. 476
Pye, C. 426, 430
Pykkänen, L. 102, 107, 260, 274, 275, 342
Pyykkönen, P. 116

Qu, A. 638n

Rabin, J. 431
Rackowski, A. 188
Radhakrishnan, R. 526, 527, 528, 529
Rainer, F. 18, 27, 70, 241, 243, 249, 338, 349, 351, 352, 386n, 410, 750
Ralli, A. 308
Ramat, P. 454, 772–3
Ramchand, G. C. 261
Ramscar, M. 101, 102, 105, 110, 114, 115, 117
Ramstedt, G. J. 493
Ranjan, R. 463n
Rankin, J. 426, 431
Rappaport, M. 277
Rappaport Hovav, M. 55, 90, 160n, 239, 240, 244–5, 248, 253–4, 259, 261, 262, 266, 267, 401
Rastle, K. 91, 96, 99, 100, 104
Rau, D.-H. V. 142n
Rayner, K. 97, 99, 108
Read, B. 101
Refsing, K. 312–13
Regamey, C. 624
Regier, T. 180–1, 184
Reh, M. 601, 771
Reichard, G. 174
Reid, L. A. 182, 521
Reinisch, L. 577
Reintges, C. 575, 580
Rescorla, R. 114
Revel-Macdonald, N. 556
Rice, K. 4, 132, 187, 371, 373, 678–9, 682–4, 687–9, 692–3, 742, 784
Riddle, E. 72

Riese, T. 475, 483
Rimzhim, A. 108
Risch, E. 14
Rischel, J. 526, 709
Robbeets, M. 493
Roberts, I. 260
Roberts, J. 182
Robin, F. 622
Robins, R. H. 24, 138
Robinson, J. O. 173
Robson, J. 101
Robustelli, C. 119
Rodrigues, A. S. 450
Roelofs, A. 111, 112, 225
Roeper, T. 50, 239, 246, 248, 252, 254
Róna-Tas, A. 630
Rosch, E. 12
Rosen, N. 48, 172
Rösler, F. 104
Ross, J. R. 236
Ross, M. D. 768
Ross, S. E. 109
Rossdeutscher, A. 240
Rothmayr, A. 286
Rothstein, S. 259, 276
Roy, C. 431
Roy, I. 255
Rubin, G. 99
Rubino, C. 145, 169, 174, 175, 176, 549
Rueckl, J. G. 108
Ruhl, C. 340
Rumelhart, D. E. 103
Ruong, I. 475, 481
Russell, K. 48, 172
Ryan, K. M. 188, 372, 376–7, 378, 380
Ryder, M. E. 238, 248, 249, 422–3

Saciuk, B. 374
Sadock, J. 705, 714, 716, 719
Saeed, J. I. 162n, 581
Sag, I. A. 12, 260, 262
Sagart, L. 622, 624, 625, 626, 629, 630
Sagi, Y. 433, 436
Säily, T. 408–9
Sakel, J. 311
Samek-Lodovici, V. 209
Sammallahti, P. 481
Sanchez, T. S. 151

- Sander, P. 101
 Sanders, G. 155, 157, 160n, 161
 Sanders, L. 101
 Sandra, D. 104
 Sapir, D. J. 356
 Sapir, E. 162, 297, 725, 728, 729n, 783
 Sauer, G. 475
 Saussure, F. de 51–2, 54, 355
 Savà, G. 582
 Savickienė, I. 308
 Saxon, L. 684
 Scalise, S. 15, 18, 19, 20, 27, 50, 293, 298–9, 303, 357
 Scarborough, R. A. 112
 Scarpa, G. F. 745, 750, 753, 755–6, 758–9
 Schachter, P. 174
 Schadeberg, T. C. 560n
 Schaeffer, R. P. 649n
 Schaerlakens, A. M. 436
 Schäfer, F. 238–9, 240, 241, 244, 246–7, 248, 262n, 265, 267
 Schiller, N. O. 111, 112
 Schilling, G. 13
 Schmid, H.-J. 34, 409
 Schmidt, W. 520–1
 Schneider, E. W. 339
 Schneider, K. 308
 Schneider-Blum, G. 600
 Schreuder, R. 32, 91, 97, 101, 104, 105, 106, 107, 108, 109, 110, 111–12, 114, 116, 225, 341
 Schriefers, H. 99, 101
 Schröder, A. 395
 Schuessler, A. 625
 Schuh, R. 584, 585, 587
 Schultink, H. 427, 433, 438
 Schultze-Berndt, E. 656
 Schuppler, B. 109, 112
 Schürmann, C. W. 653, 654n, 657
 Schvaneveldt, R. 437
 Scorza, D. P. 148
 Segalowitz, S. 99
 Segui, J. 109, 430, 439
 Seidenberg, M. 114, 115
 Seidler, S. 434, 436
 Selkirk, E. O. 6, 42, 86, 131, 227, 374, 375
 Sells, P. 260
 Serbat, G. 344
 Sergeant-Marshall, S. 96
 Shafer, R. 629
 Shaoul, C. 101, 105, 115, 117
 Shapiro, M. C. 470
 Sharp, J. 660–2
 Shay, E. 574, 577, 579, 581, 583, 584, 585, 586, 587–8, 589
 Shibatani, M. 124
 Shidanluo 631
 Shiota, J. 92
 Short, D. 446
 Shorto, H. L. 150, 526
 Shtyrov, Y. 108
 Sibanda, G. 186
 Sidwell, P. 150, 493, 520, 527, 530
 Siegel, D. 6, 36, 93, 131, 374–5
 Siirinen, M. 478
 Silfverberg, L. 476
 Silnitsky, G. G. 265
 Siloni, T. 267n
 Silver, S. 725
 Simon, W. 624, 625n, 628
 Simons, G. F. 670
 Simpson, G. B. 96, 99
 Simpson, J. 652, 654, 656, 659, 664, 742
 Sims, A. 50, 338, 378
 Singh, R. 171–2, 458, 462n, 467
 Siptár, P. 486
 Skousen, R. 395
 Slater, K. W. 495
 Slobin, D. I. 425
 Sloutsky, V. M. 430, 434, 439
 Smet, H. De 156n
 Smith, C. 266, 690
 Smith, K. D. 526
 Smith, L. R. 704, 707, 710, 718, 720
 Smith, R. L. 543
 Smolensky, P. 141
 Smolka, E. 104
 Snellgrove, D. 630
 Söderbergh, R. 436
 Soenjono Darjowidjojo 548
 Solms, H.-J. 386
 Solomon, N. P. 112
 Solomyak, O. 92, 102, 104
 Sommer, W. 99
 Song, J. J. 777, 783
 Spalding, A. 715, 717
 Spalding, T. 32

- Spencer, A. 7, 18, 50, 63, 93, 156, 163, 172, 204, 209, 210, 215, 216, 258, 375, 726
- Spieler, D. 96, 99
- Spinelli, E. 113
- Splett, J. 386
- Spreng, B. 718
- Sproat, R. 64, 65n, 105
- Stamatakis, E. A. 104
- Stampe, D. 521
- Stankiewicz, E. 298
- Starosta, S. 541
- Starostin, S. 493
- Steele, J. S. 53, 666
- Steels, L. 80
- Stein, A. 396, 397, 403–5, 436
- Steinhauer, K. 101
- Štekauer, P. 4, 6, 7–8, 18, 50, 121, 140–1, 146n, 155, 157, 165, 175, 237, 242, 282, 293n, 305, 309, 315–16, 322n, 327n, 357, 358, 373–4, 381, 410, 412–19, 422, 448, 449, 452, 453, 463, 779, 780, 782–4, 786
- Stell, N. 745, 747, 751–2, 754–6, 758–60, 762–4
- Stephany, U. 426
- Steriade, D. 170
- Sterling, C. 431
- Stern, C. 436
- Stern, W. 436
- Stevens, A. M. 550
- Stockwell, R. 221
- Stolz, C. 769
- Stolz, T. 170, 174, 176, 182, 184, 492, 769
- Stonham, J. 177–8
- Storch, A. 603n
- Stowell, T. 277n
- Street, J. C. 493
- Stroh, C. 170, 174, 176, 182, 184, 492
- Stump, G. 6, 18, 52–4, 64, 113, 140, 260, 299–300, 334, 335, 354, 356, 358–9, 361, 365–6, 742, 757
- Sugioka, Y. 92, 293
- Sun, J. T.-S. 631, 635, 636, 646, 649n, 786
- Sundby, B. 160n
- Sunik, O. 495
- Surabhi, B. 468
- Sutton, P. 654
- Suzuki, R. 308
- Suzuki, S. 395
- Svaib, T. 426, 431
- Svantesson, J.-O. 528, 530, 531, 532, 533, 535
- Svenonius, P. 190, 191–2, 205, 207n, 210, 213–14, 217
- Swadesh, M. 178, 743
- Swan, O. 224
- Sweet, H. 157, 158
- Szymanek, B. 35, 297, 308, 356, 757
- Tạ Văn Thông 526, 529, 530, 531, 532, 538, 539, 540, 541, 543
- Tabak, W. 101, 111–12
- Tacconi, T. 745, 760
- Tadmore, U. 610
- Taft, M. 91, 103–4
- Taine-Cheikh, C. 579, 581
- Talamo, L. 378
- Tamanji, P. N. 309
- Tan, K. 638n
- Tanguy, L. 81
- Tannenbaum, M. 108
- Taran, V. L. 427
- Tarasova, L. 122
- Tarp, S. 339
- Taylor, J. R. 181
- Teichelmann, C. G. 653, 654n, 657
- Teleman, U. 332
- Temmerman, R. 12
- ten Hacken, P. 12, 18, 19, 23, 24, 41, 43, 357
- Tenishev, E. R. 495
- Terasawa, J. 395
- Tersis, N. 703
- Thieberger, N. 554
- Thomas, D. 529, 530, 533, 535, 543
- Thomason, S. G. 151–2, 768, 774, 776
- Thompson, C. 683
- Thompson, L. 531, 535, 538, 541, 542
- Thompson, P. 396
- Thompson, S. A. 237
- Thornton, A. M. 105, 106
- Thráinsson, H. 447–8
- Thurgood, G. 153, 609
- Timberlake, A. 226
- Tinbergen, D. C. 436
- Tolskaya, M. 495
- Tomaščíková, S. 410, 412–19
- Tompa, J. 486
- Tompo ole Mpaayei, J. 600, 607
- Toosarvandani, M. 726

- Topping, D. M. 144n, 556
 Torre Alonso, R. 396
 Tournadre, N. 621
 Tournier, J. 154, 158
 Tovar, A. 743, 744
 Toweett, T. 594, 595
 Townsend, C. E. 447
 Trager, G. L. 14–15
 Traugott, E. C. 319, 349
 Treiman, R. 96, 99
 Tremblay, A. 112, 113
 Trips, C. 30, 236, 396, 397, 399–402, 403–5
 Trnka, B. 156n, 161n, 162n, 163n
 Trnková, J. 308
 Trommer, J. 205, 207, 214
 Trubetzkoy, N. 298
 Trussel, S. 149, 150
 Tsujimura, N. 192n, 201, 217, 778
 Tucker, A. N. 594, 600, 601, 604, 607
 Tucker, B. V. 112
 Tuggy, D. 50
 Turano, K. 99
 Turon, L. 296–7, 308
 Tyler, A. 431, 433
 Tyler, L. K. 104

 Ubrjatova, E. I. 495
 Uebach, H. 624, 625
 Ueda, I. 207–9
 Uhlenbeck, E. M. 113, 380
 Ullman, M. 101, 102
 Ultan, R. 6, 140, 146, 147, 148, 150, 315
 Unseth, P. 594
 Uray, G. 622, 627, 628
 Urbanczyk, S. 171, 176, 185, 308
 Urdze, A. 170, 174, 176, 182, 184, 492
 Ushakova, T. N. 436
 Ussishkin, A. 108, 115, 196n, 214

 Vajda, E. J. 509, 510n, 511–12, 513n, 517, 518, 784, 786
 Valera, S. 4, 7–8, 140–1, 146n, 237, 242, 305, 309, 315–16, 322n, 327n, 373–4, 381, 448, 449, 452, 453, 463, 782–4, 786
 van den Berg, R. 312, 549
 van Der Lugt, A. 379
 Van der Voort, H. 183
 van Dommelen, W. A. 109, 112
 van Driem, G. 609, 621, 630

 Van Dyke, J. 101, 106
 Van Geenhoven, V. 714, 716
 van Hout, A. 239, 246, 248, 252, 254
 Van Hout, R. 110, 151, 410, 411–12
 van Marle, J. 6, 70, 154, 155, 157, 162n, 163, 228, 354, 356, 360, 362, 364–5
 Van Valin, R. D. 259, 261
 van Voorst, J. 244
 Vannest, J. 116
 Varela, S. 27
 Vasilevič, G. 495
 Vaux, B. 171, 231
 Vázquez-López, V. 396
 Velupillai, V. 777
 Veneziano, E. 426
 Verma, M. K. 467
 Veselinova, L. 637
 Vidal, A. 744–5, 748, 754, 764
 Viegas Barros, P. 745
 Viên Ngôn Ngữ Học 535
 Vilkuna, M. 476, 478, 480
 Vinokurova, N. 235, 246, 248
 Virkkala, M. M. 436
 Vitevitch, M. S. 112
 Vittrant, A. 622
 Voeltz, E. 560
 Vogel, P. M. 317
 Volkart, M. 623
 von Fintel, K. 274
 von Schroeder 41
 Vorbichler, A. 599
 Vossen, R. 603n
 Vovin, A. 494, 495
 Vu, S. 537, 540
 Vul, E. 98
 Vycichl, W. 575, 578

 Waag, C. 600, 601, 604
 Wade, T. 226
 Wagner, A. 114
 Wagner-Nagy, B. B. 475, 489
 Walker, A. 651
 Walker, J. 93
 Walker, R. 215
 Walter, M. 625n
 Wasow, T. 277
 Waszakowa, K. 356
 Watson, J. 308

- Watson, R. 522, 533, 535, 537
Watson, S. 530, 533, 542
Watts, J. 149
Weingarten, R. 112–13
Weinreich, U. 151, 768
Weiss, D. 599
Wells, R. 68
Wermser, R. 393
Werner, H. 509, 510–11, 514
Wetzer, H. 258
Whaley, L. J. 777
Wheeler, M. W. 454–6
Whitney, W. D. 151, 323, 327, 328–31, 335
Whorf, B. L. 156n, 725
Wiebusch, T. 610
Wiedenhof, J. 25
Wierzbicka, A. 55, 303–4, 757
Wilhelm, A. 684
Wilkins, D. P. 656, 657, 666–7
Will, U. 112–13
Williams, E. 6, 86, 133, 246, 358
Williams-van Klinken, C. 144n
Wilmanns, W. 41
Wilson, J. L. 148
Wilson, S. 656
Wiltshire, C. 47
Windsor, J. 427, 431, 433, 437
Winkelman, P. 98
Winters, K. E. 721
Wise, H. 450, 454–6
Withgott, M. 742
Wittgenstein, L. 341
Wolf, M. 192, 209, 210, 213, 231–2
Wolfenden, S. N. 629, 630
Wolff, J. U. 546
Wolters, G. 114
Wood, E. J. 138, 140, 173, 350
Wood, S. N. 106
Woodbury, A. 703, 704n, 706–7, 712, 715n, 719, 722
Wright, S. 266
Wunderlich, D. 262
Wurm, L. H. 106, 107, 109, 110, 116
Wysocki, K. 431
Yallop, C. 652
Yamanouchi, K. 395
Yang, C. 69n
Yap, M. 96, 99
Yarlett, D. 102
Yeh, M. 241
Yeung, N. 115
Young, R. W. 671, 675, 681, 690, 699–700
Yu, A. 6, 138, 140, 141, 143, 145, 146, 147–8, 188, 205, 216, 227
Zadoks, A. 621
Zajceva, M. I. 475
Zandvoort, R. W. 155, 158
Zawaydeh, B. A. 196–7
Zé Amvela, E. 159n
Zeisler, B. 624, 625
Zeller, J. 315
Zemskaja, E. A. 361
Zepeda, O. 205–6
Zergollern-Miletic, L. 395
Zevenbergen, A. 436
Zhang, H. 610, 613
Zhang, L. 622
Zheng, X. 99
Zide, N. 151, 521
Zigmond, M. L. 729n
Zimmer, K. E. 410
Zimmermann, E. 205, 207, 214
Zimmermann, J. 316
Zirkel, L. 378
Zöfgen, E. 339
Zoll, C. 6, 61, 169–70, 171, 177, 179, 182, 187, 222, 556
Zorc, R. D. 651, 657
Zwanenburg, W. 35, 386n
Zwicky, A. 18, 293, 296, 321n
Zwitserslood, P. 99, 101, 111

SUBJECT INDEX

- a-morphous morphology 16
- a-templatic morphology 191, 205–17
- ablauts 387 *see also* apophony
- acyclicity of affix-ordering 377–80
- adjectival derivation 276–95
 - Afroasiatic languages 581–2
 - Altaic languages 497–8
 - deadjectival adjectives 280, 364–5
 - denominal adjectives 279–80
 - affixation 454, 466
 - conversion 451
 - relational vs. qualitative 284–5
 - systematic properties 281–3
 - deverbal adjectives 259n, 276–8, 281
 - affixation 454, 466
 - conversion 451
 - open vs. closed scale 286
 - Finnish 280, 477–8
 - Hungarian 486–7
 - Mari 483
 - Mordvin 481–2
 - Nenets 490
 - Permic languages 485
 - Yeniseian languages 511–12
- adjectives 24, 330
 - affixation, systematic properties 281–3
 - of degree, nominalization 287–8
 - gradability 284–6
 - as lexical category 290
 - post-nominal, Athabaskan languages 683
 - proprietary 503–8
 - reduplication in Mandarin Chinese 615
- adpositional derivation, Nilo-Saharan languages 603
- adverbial derivation 479
 - Afroasiatic languages 582–3
 - Athabaskan languages 673
 - deadjectival adverbs
 - affixation 455
 - conversion 451
 - non-predicative 294–5
 - predicative 291–3
 - denominal adverbs 294
 - Hungarian 488
 - Mari 483–4
 - Mataguan languages 747–8
 - Mon-Khmer languages 538–9
 - Nenets 491
 - Yeniseian languages 512
- adverbial morphology 293
- adverbs 24, 329
 - ideophonic 603–4
 - indefinite degree vs. modifiers of
 - extension 284–5
 - non-predicative 294–5
 - predicative 291–3
 - reduplication in Mandarin Chinese 615
- affixation 16–17, 118–19
 - allomorphs 125–6, 219–20
 - autosegmental 193, 209–17
 - vs. back-formation 45
 - vs. compounding 26–7, 48–9
 - bound lexemes 34–6
 - conversion 43–5
 - incorporating languages 38–41
 - prefixation 36–7
 - semi-suffixes 30–2, 33–4
 - synthetic compounds 41–3, 45
 - transition 28–30, 36
 - concatenative 200–4
 - continuous 124–6
 - discontinuous 126–8
 - evaluative morphology *see* evaluative morphology (EM)
 - frequencies 7–8

affixation (*Cont.*)

- function words 324, 327–8
- historical change 384–406
- in language acquisition 424–8, 430–7, 439
- languages and language families
 - Altaic languages 496–501
 - Athabaskan languages 672–84, 685–700
 - Austronesian languages 546–50
 - Chadic languages 583–7
 - Eskimo-Aleut languages 704–10
 - Indo-European (IE) languages 446, 451–7
 - Indo-Aryan languages 460–72
 - Inuit 712–22
 - Mandarin Chinese 611, 613, 616–20, 625–6
 - Mataguan languages 745–66
 - Mon-Khmer languages 523–34
 - Niger-Congo languages 559–72
 - Nilo-Saharan languages 593–608
 - Pama-Nyungan languages 658–68
 - Rgyalrong 634, 637–49
 - Tibetan 623–30
 - Uralic languages 475–92
 - Uto-Aztecan languages 729–42
- absence in Vietnamese 7
- Yeniseian languages 510–19
- in neoclassical word formation 120
- lexical affixes 38–41
- meanings 134–5, 144–6
- nominalization *see* nominalization
- novel forms 90–1
- numerals 335
- ordering 93–4, 130–2, 370–1
 - acyclicity 377–80
 - co-occurrence 371–2
 - derivation vs. inflection 371
 - free variation 376–7
 - multiple strata models 374–6, 378
 - reduplication 185–9, 380–2
 - relative 19
 - semantic factors 372–4
- polysemy 343–9
 - of English *-er* suffix 342–3
- position
 - circumfixation 8, 127–8
 - effect on processing speed 128–30
 - infixation 8, 125–6, 136–9, 140–3 with
 - reduplication 537, 555–6
 - borrowed 151–3
 - derivational 139–40
 - meanings 144–6
 - stability over time 149–51
 - universals 146–8
 - interfixation 126, 146n
 - prefixation 8, 20, 36–7, 133
 - instrumental prefixes 727–9
 - suffixation 7–8
 - appreciative suffixes 296–7
 - borrowed suffixes 76–80
 - ordering 116
- prepositions as 36–7
- processes as 122–4, 139
- processing studies 104–8
- production studies 112–13
- productivity *see* productivity
- propriative adjectives 504–8
- reduplication 177–8, 182–4 *see also*
 - reduplication
- relational vs. qualitative adjectives 285
- replication, grammatical 767–76
- rival affixes/suppletive allomorphy 70–2,
 - 74–5, 80–3, 220–9
- root-and-pattern morphology 193–5,
 - 575–6, 579
- semantic change 350–2
- semi-affixes/affixoids 30–2, 121–2
- socio-geographic variation 411–12
- sociolinguistic factors 408–9
- systematic properties 281–3
- unique affixes 121, 131
- universals 146–8, 783, 785–6
- verbal derivation 87, 257–65
- voice 546–7
- zero-form affixes *see* conversion
- age differences 101
 - in second language use 420
- agentives 453, 466–8 (*see also* nominalization)
 - in language acquisition 431–6
- agglutinating morphology 725–6, 740–2
- agreement 22–3, 25
- allomorphy
 - derivation vs. inflection 227–9
 - phonologically vs. lexically
 - conditioned 125–6
 - suppletive 70–2, 74–5, 80–3, 220–30
 - directionality 231–4
 - types 219–20, 225

- alternation *see* apophony
 ambiguity *see* polysemy
 analogical derivations 45–6, 47
 ‘ancient’ languages 148–51
 antipassives (*see also* passives)
 Bantu languages 562
 Niger-Congo languages 566
 Rgyalrong 639–40
 aphasia, studies 92
 applicatives 274–5, 638, 733–6, 764–5
 suffixation 219, 223–4, 560, 563–4, 737–41
 apophony 8, 123–4
 ablauts 387
 Nilo-Saharan languages 598
 Pama-Nyungan languages 665
 Rgyalrong 635–7
 umlauts 212–13
 Articulated Morphology 53
 aspect-marking, reduplication as 173–4
 aspectuals
 Athabaskan languages 693–4
 Mon-Khmer languages 531
 auditory comprehension studies 108–10
 augmentatives *see* evaluative morphology (EM)
 autosegmental affixation 193, 209–17
 auxiliary verbs *see* function words
- back-formation 45, 779
 backwards word lists 87
 banks of units 114
 base components 15
 base marking, function words 322, 323
 blends 46–7
 blocking 70–2, 86–7, 228, 229
 body morphemes 16
 borrowing
 compounds 121–2
 Greek affixes in English 767
 infixes 151–3
 Mandarin Chinese 610
 Pama-Nyungan languages 653–5, 657, 661–3
 structural 767–76
 bound morphemes 64
 bound roots 34–6, 38, 40
 boundedness 285
- British National Corpus (BNC) 90
see also corpus data
- C1-LIN (cyclicity-linearity) model 232
 calques 768
see also borrowing
 case marking
 Nilo-Saharan languages 606–8
 Polish 10–11
 reduplication as 173
 categories, mixed 503
 categorizing tradition 13–14
 category changing 4, 5, 11, 15, 19–20, 86–7 (*see also* adjectival derivation; adverbial derivation; conversion; nominalization; verbal derivation)
 adjectives and adverbs 24
 affixation *see* affixation
 by apophony 123–4
 Eurocentric view 7
 reduplication as 174–5
 by stress alternation *see* conversion
 Category-Conditioned Degree of Productivity 73
 category maintaining morphology 456, 464
 Altaic languages 498–501
 Athabaskan languages 679–84
 Nilo-Saharan languages 599–601, 606
 Pama-Nyungan languages 664–8
 Rgyalrong 637–42
 verbs 479
 causatives 87
 contact-induced grammaticalization 773–4
 inchoative alternation 265–73
 Mon-Khmer languages 526–7
 Niger-Congo languages 564
 Rgyalrong 638–9
 Tibetan 623–4
 Uto-Aztecan languages 739
 change of state verbs *see* verbal derivation
 circumfixation *see* affixation
 clipping 385
 clitics 23
 Cognitive Linguistics 340
 coinage, conversion as 156
 coindexing 56–7
 comparatives 203, 284, 551
 and superlatives 72, 202, 286–9, 332

- completive/incompletive forms 214
- complex signs
 semantic representation 52, 58–61
 sensory-motor representation 61–3
- complex uniqueness points (CUP) 109
- complex words
 lexical decision tasks 95–7
 processing 103–10, 115–16
 production studies 111–13
- Complexity Based Ordering 93, 378–80
 see also affixation, ordering
- compounding
 vs. affixation *see* affixation, vs.
 compounding
 Afroasiatic languages 578
 Altaic languages 502
 ambiguous novel compounds 32–3
 analogical derivations 45–6, 47
 Austronesian languages 551–2
 blends 46–7
 borrowed words 121–2
 definition 3
 vs. derivation 4, 6, 26
 in evaluative morphology (EM) 312, 314
 function words 319–20
 headedness 133
 Indo-European (IE) languages 446, 447–9
 interfixes 126
 in language acquisition 427, 431, 433
 Mandarin Chinese 610, 611, 316
 Mon-Khmer languages 523, 540, 541–3
 neoclassical word formation 35–6, 45,
 119–20, 134, 385
 Pama-Nyungan languages 655–6
 reduplication 47–8, 459
 replication in Finnish 769–70
 sociolinguistic factors 422
 Tibetan 623, 625
- conceptual representations 54–61
- Conceptual Structure *see* encyclopedic knowledge
- consonant alternation *see* apophony
- constants *see* encyclopedic knowledge
- Construction Grammar 397
- Construction Morphology 17, 283, 347–9
- constructionist theory 282–3
- contact-induced grammaticalization 403–6,
 769–76 *see also* borrowing
- context 114
- conversion 7, 8, 43–5120–1, 133, 161–4, 356
 (*see also* category changing)
 Altaic languages 501
 cross-linguistic problems 158, 164
 cross-linguistic tests 164–7
 as derivation 157–9
 Finnish 479–80
 formal identity 159–61
 function words 321–2
 Hungarian 488–9
 Indo-Aryan languages 460
 Indo-European (IE) languages 446, 449–51
 in language acquisition 428–30
 lexical derivation 154–7
 Mandarin Chinese 620
 Mari 484
 Nenets 491
 stress alternation 122–3, 159, 216
 Uto-Aztecan languages 731
 verbal derivation 258–60
- Corpus of Contemporary American English
 (COCA) 86, 90
- corpus data 85–7, 89–91
 convergence with psycholinguistic
 study 92–4
 hapaxes 73–6
- cross-linguistic variation, diminutives 25
- data collection 85–7, 88–91, 777–8
- deadjectival adjectives *see* adjectival
 derivation
- deadjectival adverbs *see* adverbial derivation
- deadjectival nouns *see* nominalization
- deadjectival verbs *see* verbal derivation
- de-adverbial nouns, conversion 451
- decompositional models of speech
 production 111–12, 113, 115
- Deep Structure 15
- deletion 205–7
- denominal adjectives *see* adjectival
 derivation
- denominal adverbs *see* adverbial derivation
- denominal heads 43
- denominal nouns 246, 248, 578–9 *see also*
 nominalization
- denominal verbs *see* verbal derivation
- deprepositional verbs/nouns, conversion 451

- derivation
 vs. compounding 4, 6, 26
 concatenative 61–2, 118–35, 200–4, 727, 785–6
 conversion as 157–9
 definition 3–4
 historical change 384–406
 infixation as 139–40
 vs. inflection 6, 10–11, 17–18
 adverbial morphology 293
 affix-ordering 371
 allomorphy 227–9
 approaches 13–17
 Athabaskan languages 680
 borderline cases 24–5
 comparatives and superlatives 287–9
 criteria 19–24
 defining boundaries 779
 evaluative morphology (EM) 298–300, 309
 infixation 146–8
 paradigms 354, 357–61
 reduplication 175–6
 verbal derivation 260
 non-concatenative 169–218, 709–10, 729, 785–6
 paradigms 356, 361–9
 reduplication as 174–5
 Derivational Paradigm Functions (DPFs) 366
 Derived Stems 740
 derived words, studies 91–2
 desideratives 532
 deverbal adjectives *see* adjectival derivation
 deverbal nouns *see* nominalization
 dictionaries 76–80, 87–8
 as analogous with mental lexicon 102–7, 110–11
 digraphs, processing 115–16
 diminutives *see* evaluative morphology (EM)
 directionality of conversion 160–1
 discovery procedures 84
 discreteness of language 67–70
 distributed collectionist models 114
 Distributed Morphology (DM) 16, 53, 232, 246, 263–4
 dual route models 106
 dvandvas 448–9
see also compounding
- Edge-Bias Effect 141
 education level, effect on second language use 420–1
 EEG (electroencephalography) 98–9, 109
 elative forms 213
 emphatics 207–9
 encyclopedic knowledge 55, 57–9, 113
 ETM *see* Google Search
 evaluative morphology (EM) 4–5, 25, 207–9, 296–7, 300–2, 456–7
 Afroasiatic languages 578–9
 Altaic languages 498–9
 Athabaskan languages 681–4
 compounding 312, 314
 defining boundaries 779–80
 derivation vs. inflection 298–300, 309
 Finnish 477
 language acquisition, diminutives in 426, 430–1
 Mataguan languages 757–9, 761
 metaphors for diminutives 301–2
 onomasiological theory 302–3
 Pama-Nyungan languages 667–8
 phonological iconicity 314–16
 polysemy 349–50
 proprietive adjectives 508
 quantitative vs. qualitative 303–8
 reduplication 181, 311–12, 314
 semantics 300–2, 304–8
 sociolinguistic factors 407–8, 420n
 Tibetan 627
 world languages 309–14
 Evaluative Morphology Saturation 307–8
 ‘expressive morphology’ 296–7, 298
 extragrammatical morphology 296
 eye-tracking 116
 methodologies 97, 106
- f-morphemes 16 *see also* Distributed Morphology (DM)
f-structure 12
 fMRI (functional magnetic resonance imaging) 98, 99, 104
 frequency 289
 function words 317–19
 derivation 319–32, 337
 functions, in semantic skeletons 56–61

- gender, nominal 465
gender differences 101
 in affix use 408–9
 in second language use 420
Generalized Template Theory (GTT) 185
generative grammar 15
generative movement/tradition 5, 84
Generative Semantics 15
Google Books 90
Google Search 80–3
gradability of adjectives 284–6
grammars of languages 14
grammaticalization
 bound roots to lexical affixes 40
 contact-induced 769–76
 of function words 319–20, 323
 lexemes to affixes 28–30, 40

habilitative forms 198–9
handbooks on morphology 18
hapaxes 73–6, 404, 411
head position 129
 in compounding 27
Head-Driven Phrase structure Grammar 12
headedness 132–4
homonymy 339, 351–2 *see also* polysemy
homophony, Athabaskan languages 699–700
hypocoristic forms 196–7, 201–2

iconicity, in evaluative morphology 314–16
ideophones 7
 adverbs 603–4
 Afroasiatic languages 581
 Chadic languages 589
 Rgyalrong 646–9
idiosyncratic speech 89
imperfective form 209
implicit priming paradigm 111
inceptives 694–5
inchoative/causative alternation 265–73, 580
 see also verbal derivation
incorporation 38, 40, 470, 513, 518, 607–8, 620
 Athabaskan languages 699
 Eskimo-Aleut languages 714, 716
 Mataguayan languages 746
 Uto-Aztecan languages 728, 732
incremental mapping systems 53

inferential theories of morphology 52–3, 60–1
infinitives, Yeniseian languages 518–19
infixation *see* affixation
inflection 4–6
 absence in some languages 7
 vs. derivation *see* derivation, vs. inflection
 mapping 53–4
 paradigms 354–61
 reduplication as 173–4
inflectional entropy 111–12
Inflectional Paradigm Functions (IPFs) 366
Inner Stems 740
inseparable complex verbs (ICVs) 397–8
instrumentals 128, 153, 240–9, 277, 352, 426,
 529, 753–5
 prefixation 727–9
 suffixation 431–4
intercategorical relationships 365
involuntary prefixes, Mon-Khmer
 languages 533
irregular forms 68–9

kinship terms
 Athabaskan languages 673
 Pama-Nyungan languages 666–7

l-morphemes 16 *see also* Distributed
 Morphology (DM)
Lambda-Abstraction Specification 302
language acquisition 424–39
language background, effect on second
 language use 421–2
language change, affixation 384–406
language contact, grammaticalization 403–6,
 769–76
learning models 114
legal concepts, terminological definitions 12
letter transpositions in processing studies 108
lexeme formation *see* neologisms
lexemes 17, 118–19
 transition to suffixes 28–30
lexical affixes 38–41
lexical allomorphy 224–5
lexical decision experiments 91–2, 95–7,
 116, 117
lexical decomposition 261–3, 265
lexical derivation, conversion as 154–7

- Lexical-Functional Grammar 12
 Lexical Morphology and Phonology (LMP) 93
 lexical redundancy rules 17
 lexical semantics 16
 lexical stock expansion 7–8, 16
 lexical theories of morphology 52–3, 60, 63–4
 listening, processing studies 108–10
 loanwords *see* borrowing
 locatives 242, 248–9, 343–4, 748
 affixation 350, 352, 393
 Mataguyan languages 755–7, 765
 Mon-Khmer languages 529, 533
 Turkic languages 497–8
 Yeniseian languages 513
 mapping 52–4, 63–5
 MEG (magnetoencephalography) 98, 99
 mental lexicon 33–4, 35
 organization 102–7, 110–11
 methodology
 convergence 92–4
 data sources 85–7, 88–91
 experimental methods 95–9, 101–2
 issues in finding universals 777–80
 materials 99–100
 morphological theory 100–2
 production studies 110–13
 in psycholinguistic study 91–4
 word processing studies 103–10
 model languages *see* replication
 modifiers of extension 284–5
 morpheme frequency effects 115
 Morpheme-to-Seme-Assignment Principle (MSAP) 414
 morphemes
 role in morphology 16
 Saussure's treatment of 52n
 suppletive allomorphy 229–30
 morphological categories 364–5
 Morphological Doubling Theory 171
 morphological family size effect 107, 109, 115
 morphological phrases 43
 morphological theories
 lexical vs. inferential 52–3
 methodologies 100–2
 morphological transparency 33
 morphology
 augmentative 207–9
 challenge to discreteness theories 68
 subtractive 205–7, 550
 morphemes, suppletive allomorphy 230
 morphosyntactic allomorphy 223–4, 226–7, 230
 naïve discrimination learning (NDL)
 model 114–16
 natural concepts 11–12
 negation 125, 328, 490–1
 Altaic languages 500
 Eskimo-Aleut languages 708, 718, 722
 in language acquisition 434–5
 Mon-Khmer languages 532–3
 Pama-Nyungan languages 666
 particles 488
 prefixation 27, 30, 86, 134, 388, 393, 426–7, 475
 suffixation 478
 neo-Saussurrean sign 53, 54–61, 65
 neoclassical compounds 35–6, 45, 119–20, 134, 385
 neologisms 16, 37–8, 360–1 (*see also* language acquisition)
 blocking in 70–2
 in data collection 85–6, 88–9, 90
 defining boundaries 779
 Early Modern English (EmodE) 393
 hapaxes as 74
 sociolinguistic factors 409, 410–11, 412–19, 422–3
 neuroimaging techniques 98–9, 104, 108, 109
 nominalization 15, 86–7, 235–6, 453–4, 464–70 (*see also* category changing)
 action/state 237
 deadjectival nouns 248, 254–6, 287–8, 579
 affixation 464
 conversion 450–1
 de-adverbial nouns 451
 denominal nouns 246, 248, 578–9
 deprepositional nouns 451
 deverbal nouns 42–3, 44, 160–1, 237–54
 affixation 464–6
 conversion 450, 460
 event/result 253–4

nominalization (*Cont.*)

languages and language families

Afroasiatic languages 576–9

Altaic languages 497

Athabaskan languages 672–3, 674–6

Chadic languages 584–7

Chorote languages 749–50, 753, 755–6, 759

Cushitic languages 577–8, 579

Dutch 243–4, 250

English 235–40, 244–5, 249–52, 256, 453

Finnish 476–7

French 241, 255, 454

German 238, 242, 244, 255, 453–4

Greek 240–1, 242, 250, 255, 453

Hungarian 486

Inuit 712–14

Italian 453

Mari 483

Mataguan languages 748–61

Mon-Khmer languages 527–9

Mordvin 481

Nenets 489–90

Niger-Congo languages 571–2

Nilo-Saharan languages 597, 601–3, 605–6

Omotic languages 577

Pama-Nyungan languages 658–9

Permic languages 484–5

Rgyalrong 633, 643–5

Romance languages 241, 242, 244

Romanian 255

Semitic languages 578

Spanish 287–8, 453

Tibetan 624–8

Yeniseian languages 513–14

participant 237–52

polysemy 243–9

non-concatenative morphology 190–218

non-predicative deadjectival adverbs 294–5

noun class systems, Niger-Congo

languages 567–70

noun derivation *see* nominalization

noun incorporation 38, 40

noun modification, Inuit 719–21

noun stems, Rgyalrong 632

nouns

Athabaskan languages 680–1

polysemy 349–50

with proprietive adjectives 503–8

reduplication in Mandarin Chinese 614

in replication 769–70, 775–6

Yeniseian languages 510–11

number inflection in Polish 10–11

numerals, derivation 332–7

object marking, reduplication as 182

occupation, effect on second language use 421

onomasiological theories 282, 363–4, 413–22

on evaluative morphology (EM) 302–3

onomatopoeia

Chadic languages 589

Mataguan languages 746

Mon-Khmer languages 535

Pama-Nyungan languages 656–7

reduplication 458, 615

Optimal Interleaving 232

Optimality Theory (OT) 141, 151n, 192, 200, 207n, 209, 217, 231

orthographic processing studies 103–8,

112–13, 115–16

Oxford English Dictionary (OED) 76, 77n, 79, 87

paradigm of lexemes 11, 16, 21–2, 354–5

derivational 356, 361–9

inflectional 355–6

inflectional vs. derivational 354, 357–61

Parallel Architecture 16–17

parasynthesis 126–8

Partial Lexical Rules 299–300

participles 24–5

passives 24, 138, 324, 772–4

adjectival 278, 279, 281, 404–6, 478

Arabic 212n

Eskimo-Aleut languages 713, 717–18

Nilo-Saharan languages 601–2

Polish 224

Proto-Austronesian (PAN) 546

Rgyalrong 639–40

person marking, reduplication as 183

phonological alternation *see* apophony

phonological function of

reduplication 178–80

phonological iconicity in evaluative

morphology 314–16

- phonological representations 64–5
 phonologically conditioned suppletive
 allomorphy (PCSA) 221–2, 226, 228,
 230n, 231–2
 phonology (*see also* prosody)
 allomorphy 219–23, 227
 autosegmental 193, 209–17
 Mon-Khmer languages 524
 Tibetan 621–2
 picture naming tasks 97, 111–12
 pitch alternation 216–17
 plural formation 20–1, 206–7, 213–14, 231–2
 plural/pluractional reduplication 173, 178
 plurals, broken 204
 polysemy 243–9, 349–50
 affixation 343–9
 ambiguity in novel compounds 32–3
 of English *-er* suffix 342–3
 semantic change 350–2
 of verbal derivations 257–8
 possessive forms 203, 507
 reduplication as 173
 postbases, Eskimo-Aleut languages 705–6,
 711, 714–22
 potentiation of affixes 371–2 *see also*
 affixation, ordering
 pragmatic pressure in nominalization 248
 preference rules 11–12
 prefixation *see* affixation
 prepositions *see* adpositional derivation;
 function words
 presyllables 524
 pretence affixes, Mon-Khmer languages 533
 priming 92, 96–7, 98, 104, 111, 114
 process-and-paradigm model 367 *see also*
 paradigm of lexemes
 processing speed, effect of affix
 position 128–30
 productivity 288–9, 469–70
 defining 780
 diachronic studies 76–80
 discreteness 69–70
 function words 330–2
 hapaxes 73–6
 of inflectional vs. derivational
 morphology 20–2
 in language acquisition 427, 437–9
 in Mon-Khmer languages 525–32, 535
 rival suffixes 70–2, 74–5
 sociolinguistic factors 410–11, 416–19
 pronominals, Mon-Khmer languages 530–1
 proprietive adjectives 503–8
 Prosodic Morphology Hypothesis 193
 prosody 193, 200–4
 Austronesian languages 551
 conversion 122–3, 159, 216
 Uto-Aztecan languages 731–9
 proto-languages 149–51, 153, 546–7, 672
 psycholinguistics, methodology 91–4

 qualitative adjectives 284–6

 reading, processing studies 103–8, 115–16
 realizational mapping systems 53, 64
 recursion *see* reduplication
 reduplication 8, 47–8, 132, 141
 as affixation 177–8, 182–4
 in affix-ordering 185–9, 380–2
 Afroasiatic languages 579, 580
 Altaic languages 502–3
 Austronesian languages 552–7
 category-changing 174–5
 Chadic languages 587–9
 compounding 47–8, 459
 as derivation 174–5
 derivation vs. inflection 175–6
 Eskimo-Aleut languages 710
 in evaluative morphology (EM) 181, 311–12,
 314
 form and function 184–5
 iconicity 180–4
 Indo-European (IE) languages 446, 449
 Indo-Aryan languages 458–9
 inflectional 173–4
 Mandarin Chinese 614–15
 Mari 484
 Mon-Khmer languages 523, 534–41, 543
 nominalization 468–9
 Pama-Nyungan languages 656–7
 position 185–9
 as repair 178–80
 Rgyalrong 634–5
 Tibetan 622
 treatment as a form of compounding 4
 types 169–72
 universals 783, 785

- relational adjectives 284–6
- replacement approach to suppletive
 - allomorphy 229–30
- replication, grammatical 403–6, 767–76
 - see also* borrowing
- rival affixes *see* suppletive allomorphy
- root-and-pattern morphology 193–5
 - Afroasiatic languages 575–6, 579
- root/stem distinction 670–3
- rules for irregular forms 68–9

- Saussurrean sign 51–2, 54
- scalar adjectives 284–6
- scientific concepts, terminological
 - definitions 13, 14
- selectional approach to suppletive
 - allomorphy 229
- semantic body *see* encyclopedic knowledge
- semantic categories 11–12
- semantic change 350–2
- semantic decomposition 263
- semantic representations 54–61, 63
- semantic skeletons 16, 55, 56–61, 64–5, 113–14, 115–16, 245–6, 251, 347, 401
- semantic transparency effect 33
- semantic underspecification 283
- semantics
 - approach to paradigms 365–6
 - factors in affix-ordering 372–4
 - of evaluative morphology (EM) 300–2, 304–8
- semi-affixes 30–2, 121–2
- sensory-motor representation 51, 58–63
- Separation Hypothesis 24–5, 111
- sign *see also* complex signs; simplex signs
 - conceptual side 54–61
 - phonological representations 64–5
 - sensory-motor side 61–3
- simplex signs, semantic representation 54–5, 57–61
- skeletons *see* semantic skeletons
- skeptical tradition 14, 16
- sociolinguistics
 - use of complex words 422
 - system level 409–22
 - usage level 407–9
- speech production studies 110–12
 - methodologies 97
- spell-out 232
- stem-vowel alternation 387, 586–7
- stimulus-onset asynchrony (SOA) 97
- stratum models of affix-ordering 374–6, 378
 - see also* affixation, ordering
- stress alternation *see* prosody
- Strong Lexicalist Hypothesis (SLH) 15–16
- structuralist tradition 5
- subcategorization frames 63–4
- subcategorization model of suppletive
 - allomorphy 227–9, 231–2
- subject, terminological definitions 12
- subtractive morphology 205–7
 - Austronesian languages 550
- suffixation *see* affixation
- superlatives and comparatives 72, 202, 286–9, 332
- suppletion, Rgyalrong 637
- suppletive allomorphy 70–2, 74–5, 80–3, 220–30
 - directionality 231–4
- synaffixes *see* affixation, discontinuous
- synonymy avoidance 72 *see also* blocking
- syntactic categories *see* category changing
- syntactic derivation, conversion as 158
- syntactic doubling 172, 179 *see also* reduplication
- Syntactic Stems 740
- syntax, relevance in inflectional morphology 22
- synthetic compounds 41–3, 45

- templatic derivation 191, 192–204
 - Athabaskan languages 688–91
 - Yeniseian languages 514–18
- templatic meaning 262–3, 274
- tense-marking, reduplication as 174
- terminological definitions 11–12, 18, 22
 - methodological problems 778
- textbooks on morphology 17–18
- tmesis 137–8
- tonal alternation 215–16
- transposition *see* conversion
- triangle model 114
- triplication 556–7 *see also* reduplication

- umlauts 212–13 *see also* apophony
- unique bound forms 34–5
- uniqueness points (UP) 109–10
- universals 315
- approaches 780–6
- methodological issues in finding 777–80
- Urban Dictionary 58–9, 72
- ventive marking 600
- verb modification, Inuit 721–2
- verb stems, Rgyalrong 632–3
- verbal derivation 257–75
- deadjectival verbs 257, 261, 262–5, 270
- affixation 455
- conversion 451
- denominal verbs 24–5, 44, 87, 160–1, 257, 258, 350, 580
- affixation 455
- conversion 450–1, 460
- in language acquisition 428–30
- deprepositional verbs 451
- languages and language families
- Afroasiatic languages 580–1
- Altaic languages 498
- Athabaskan languages 677–9
- English 257–8, 262, 264
- Finnish 478–9
- Hungarian 350, 487–8
- Inuit 714–19
- Mari 483
- Mataguyan languages 761–6
- Mon-Khmer languages 526–7, 532
- Mordvin 482
- Nenets 490–1
- Niger-Congo languages 559–67
- Nilo-Saharan languages 603, 604–5, 606
- Pama-Nyungan languages 659–64
- Permic languages 485–6
- Rgyalrong 645–6, 648–9
- Uto-Aztecan languages 729–32, 740–2
- verbs
- Athabaskan languages 684–700
- Mon-Khmer languages, reduplication 539
- polysemy 350
- reciprocal, Mon-Khmer languages 529–30
- reduplication in Mandarin Chinese 615
- stative, in Mon-Khmer languages 530, 533
- Yeniseian languages 512–13, 514–19
- Voice Hypothesis 248
- vowel alternation *see* apophony
- vowel gradation, in Indo-European (IE) languages 447
- Weak Lexicalist Hypothesis (WLH) 15–16
- well-formedness in synthetic compounds 43
- whole-word frequency effects 104, 106, 109, 111, 115, 116
- word formation *see* neologisms
- Word Manager 18, 23
- word order, Mandarin Chinese 611–12
- word-class, problems with defining boundaries 162–4
- word-class change *see* category changing; conversion
- word-pairing studies 98
- World Atlas of Language Structures 169
- writing, production studies 112–13
- zero-derivation *see* conversion

OXFORD HANDBOOKS IN LINGUISTICS

RECENTLY PUBLISHED

THE OXFORD HANDBOOK OF APPLIED LINGUISTICS

Second edition

Edited by Robert B. Kaplan

THE OXFORD HANDBOOK OF ARABIC LINGUISTICS

Edited by Jonathan Owens

THE OXFORD HANDBOOK OF CASE

Edited by Andrej Malchukov and Andrew Spencer

THE OXFORD HANDBOOK OF COGNITIVE LINGUISTICS

Edited by Dirk Geeraerts and Hubert Cuyckens

THE OXFORD HANDBOOK OF COMPARATIVE SYNTAX

Edited by Guglielmo Cinque and Richard S. Kayne

THE OXFORD HANDBOOK OF COMPOSITIONALITY

Edited by Markus Werning, Wolfram Hinzen, and Edouard Machery

THE OXFORD HANDBOOK OF COMPOUNDING

Edited by Rochelle Lieber and Pavol Štekauer

THE OXFORD HANDBOOK OF COMPUTATIONAL LINGUISTICS

Edited by Ruslan Mitkov

THE OXFORD HANDBOOK OF CONSTRUCTION GRAMMAR

Edited by Thomas Hoffman and Graeme Trousdale

THE OXFORD HANDBOOK OF CORPUS PHONOLOGY

Edited by Jacques Durand, Ulrike Gut, and Gjert Kristoffersen

THE OXFORD HANDBOOK OF DERIVATIONAL MORPHOLOGY

Edited by Rochelle Lieber and Pavol Štekauer

THE OXFORD HANDBOOK OF GRAMMATICALIZATION

Edited by Heiko Narrog and Bernd Heine

THE OXFORD HANDBOOK OF HISTORICAL PHONOLOGY

Edited by Patrick Honeybone and Joseph Salmons

THE OXFORD HANDBOOK OF THE HISTORY OF LINGUISTICS

Edited by Keith Allan

THE OXFORD HANDBOOK OF THE HISTORY OF ENGLISH

Edited by Terttu Nevalainen and Elizabeth Closs Traugott

THE OXFORD HANDBOOK OF INFLECTION

Edited by Matthew Baerman

THE OXFORD HANDBOOK OF JAPANESE LINGUISTICS

Edited by Shigeru Miyagawa and Mamoru Saito

THE OXFORD HANDBOOK OF LABORATORY PHONOLOGY

Edited by Abigail C. Cohn, Cécile Fougeron, and Marie Hoffman

THE OXFORD HANDBOOK OF LANGUAGE EVOLUTION

Edited by Maggie Tallerman and Kathleen Gibson

THE OXFORD HANDBOOK OF LANGUAGE AND LAW

Edited by Peter Tiersma and Lawrence M. Solan

THE OXFORD HANDBOOK OF LINGUISTIC ANALYSIS

Second edition

Edited by Bernd Heine and Heiko Narrog

THE OXFORD HANDBOOK OF LINGUISTIC FIELDWORK

Edited by Nicholas Thieberger

THE OXFORD HANDBOOK OF LINGUISTIC INTERFACES

Edited by Gillian Ramchand and Charles Reiss

THE OXFORD HANDBOOK OF LINGUISTIC MINIMALISM

Edited by Cedric Boeckx

THE OXFORD HANDBOOK OF LINGUISTIC TYPOLOGY

Edited by Jae Jung Song

THE OXFORD HANDBOOK OF SOCIOLINGUISTICS

Edited by Robert Bayley, Richard Cameron, and Ceil Lucas

THE OXFORD HANDBOOK OF TENSE AND ASPECT

Edited by Robert I. Binnick

THE OXFORD HANDBOOK OF TRANSLATION STUDIES

Edited by Kirsten Malmkjaer and Kevin Windle

