978–0–19–920647–6 00-Spencer-Prelims OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) i of xxviii May 16, 2008 19:1

THE OXFORD HANDBOOK OF

CASE

THE OXFORD HANDBOOK OF

CASE

Edited by

ANDREJ MALCHUKOV

and

ANDREW SPENCER

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
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 in
Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto
With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© editorial matter and organization Andrej Malchukov and Andrew Spencer 2008
© chapters their several authors 2008
The moral rights of the authors have been asserted
Database right Oxford University Press (maker)
First published 2008
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, 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 book in any other binding or cover
and you must impose the same condition on any acquirer
British Library Cataloguing in Publication Data
Data available
Library of Congress Cataloging in Publication Data
Data available
Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
Antony Rowe Ltd., Chippenham, Wiltshire
ISBN 978–0–19–920647–6
1 3 5 7 9 10 8 6 4 2

Contents

List of Abbreviations x
Acknowledgements xvii
About the Authors xviii

Introduction 1
Andrej Malchukov and Andrew Spencer

PART I THEORETICAL APPROACHES TO CASE

1. History of the research on case 13
Barry Blake
2. Modern approaches to case: an overview 27
Miriam Butt
3. Case in GB/Minimalism 44
Jonathan David Bobaljik and Susi Wurmbrand
4. Case in Lexical-Functional Grammar 59
Miriam Butt
5. The Case Tier: a hierarchical approach to morphological case 72
Joan Maling
6. Case in Optimality Theory 88
Helen de Hoop
7. Case in Role and Reference Grammar 102
Robert D. Van Valin
8. Case in Localist Case Grammar 121
John Anderson

9. Case in Cognitive Grammar 136
Silvia Luraghi

10. Case in NSM: A reanalysis of the Polish dative 151
Anna Wierzbicka

11. Case in formal semantics 170
Helen de Hoop and Joost Zwarts

PART II MORPHOLOGY OF CASE

12. Case as a morphological phenomenon 185
Andrew Spencer

13. Case and declensional paradigms 200
James Blevins

14. Case syncretism 219
Matthew Baerman

15. The distribution of case 231
Edith A. Moravcsik

16. Asymmetry in case marking: nominal vs. pronominal systems 246
Oliver A. Iggesen

PART III SYNTAX OF CASE

17. Case, grammatical relations, and semantic roles 261
Beatrice Primus

18. Syntactic effects of morphological case 276
Ad Neeleman and Fred Weerman

19. Case and alternative strategies: word order and agreement marking 290
Anna Siewierska and Dik Bakker

20. Case marking and alignment 304
Balthasar Bickel and Johanna Nichols

21. Case and voice: case in derived constructions 322
Masayoshi Shibatani

22. Differential case marking and actancy variations 339
Andrej Malchukov and Peter de Swart

23. Case and the typology of transitivity 356
Seppo Kittilä

PART IV CASE IN (PSYCHO)LINGUISTIC DISCIPLINES

24. The acquisition of case 369
Sonia Eisenbeiss, Bhuvana Narasimhan, and Maria Voeikova

25. Case in language production 384
Alissa Melinger, Thomas Pechmann, and Sandra Pappert

26. Case in language comprehension 402
Markus Bader and Monique Lamers

27. Case in aphasia 419
Monique Lamers and Esther Ruigendijk

PART V AREAL AND DIACHRONIC ISSUES

28. Evolution of case systems 439
Leonid Kulikov

29. Grammaticalization of cases 458
Bernd Heine

30. Case in decline 470
Jóhanna Barðdal and Leonid Kulikov

31. The geography of case 479
Balthasar Bickel and Johanna Nichols

32. Case and contact linguistics 494
Lars Johanson

PART VI INDIVIDUAL CASES: CROSS-LINGUISTIC OVERVIEWS

33. Terminology of case 505
Martin Haspelmath

34. Case polysemy 518
Andrej Malchukov and Heiko Narrog

35. Marked nominatives 535
Christa König

36. Varieties of accusative 549
Seppo Kittilä and Andrej Malchukov

37. Varieties of ergative 562
Enrique Palancar

38. Varieties of dative 572
Åshild Næss

39. Varieties of genitive 581
Yury Lander

40. Varieties of instrumental 593
Heiko Narrog

41. Varieties of comitative 601
Thomas Stolz, Cornelia Stroh, and Aina Urdze

42. Spatial cases 609
Denis Creissels

43. The vocative—an outlier case 626
Michael Daniel and Andrew Spencer

44. Rare and 'exotic' cases 635
Andrej Malchukov

PART VII SKETCHES OF CASE SYSTEMS

45. Typology of case systems: parameters of variation 651
Andrej Malchukov and Andrew Spencer

46. Case marking in Daghestanian: limits of elaboration 668
Michael Daniel and Dmitry Ganenkov

47. Poor (two-term) case systems: limits of neutralization 686
Peter Arkadiev

48. Case in Iranian: from reduction and loss to innovation and renewal 700
Donald Stilo

49. From synthetic to analytic case: variation in South Slavic dialects 716
Andrej Sobolev

50. Case in an African language: Ik—how defective a case can be 730
Christa König

51. Differential case-marking of arguments in Amharic 742
Mengistu Amberber

52. Case in an Australian language: distribution of case and multiple case-marking in Nyamal 756
Alan Dench

53. Case in an Austronesian language: distinguishing case functions in Tukang Besi 770
Mark Donohue

54. Case in a topic-prominent language: pragmatic and syntactic functions of cases in Japanese 779
Akio Ogawa

55. Case in Yukaghir languages 789
Elena Maslova

56. Case relations in Tlapanec, a head-marking language 797
Søren Wichmann

57. 'Case relations' in Lao, a radically isolating language 808
Nick Enfield

References 821
Subject Index 895
Author Index 000
Language index 000

Abbreviations

*	ungrammatical
Ø	zero marker
1(sg/pl)	1st person (singular/plural)
2(sg/pl)	2nd person (singular/plural)
3(sg/pl)	3rd person (singular/plural)
A	Agent, Actor, transitive subject
ABL	ablative
ABS	absolutive
ACC	accusative
ACCin	indefinite accusative
ADEL	adelative case
ADESS	adessive
ADJ	adjective
ADJZR	adjectivizer
ADV	adverbial
AF	A(ctor)-Focus (verb form)
AFF	affirmative
AFH	Active Filler Hypothesis
AGEN	agentive
ALIEN	alienable
ALL	allative
AN	action nominal
AN	Animate
AND	andative
ANR	action nominalizer
ANT	anticipatory mood
ANTE	localization in front of the landmark
AOR	aorist
APASS	antipassive
APPL	applicative
APUD	localization near the landmark
ART	article
ASP	aspect
ASS	associative marker

ASSRT	assertive
AT	attributive form
AUG	augment
AUH	actor-undergoer hierarchy
AUX	auxiliary
AV	S(ubject)/A(actor)-voice
AVM	Attribute-Value Matrix
AY	the 'inversion' marker ay
B	bare
BEN	beneficiary
C	abstract Case
CAUS	causative
CL	clitic
CLM	clause-linkage marker
COM	comitative
COMP	complementizer
COMPL	completive
CON	construct state
CONT	localization indicating contact with the landmark, normally with the object being somehow fixed on it
COOR	Coordination marker
COP	copula(tive)
CP	Complementizer Phrase
CT	class term
CVB	converb
DAT	dative
DCT	Differential Case Theory
DECL	declarative
DEF	definite(ness marker)
DEM	demonstrative
DEM.PRON	demonstrative pronoun
DESD	desiderative
DESG	designative (case)
DET	nominal determiner
DI	the case marker di
DIR	direct case/form
DIR	directional (marker)
DIRS	directional suffix
DIS	discourse marker
DIST	distal
DISTPAST	distant past
DO	direct object

DOM	Differential Object Marking
DP	Determiner Phrase
DS	different subject marker
DSM	Differential Subject Marking
DU	dual
DUR	durative
DYN	dynamic
ECM	Exceptional Case Marking
ELA	elative
EMPH	emphatic
ENC	enclitic
EPP	Extended Projection Principle
ERG	ergative
EVID	evidential
EXCL	exclusive
EZAFE	ezafe (NP modifier connector)
F	feminine
F	Focus
F|P	Focus or P (nominal marker)
FAC	factive
FAM	familiar
FIN	finite(ness) marker
FOC	focus
FUT	future
G	given
G	goal (function/marker)
GB	Government–Binding
GEN	genitive
GER	gerund
GF	Grammatical Function
HAB	habitual
HSPM	Human Sentence Processing Mechanism
HUM	human
ILL	illative
IMP	imperative
IMPF	imperfect
IMPS	impersonal
IN	localization inside the landmark (normally, hollow container)
INAL	inalienable
INCH	inchoative
INCL	inclusive
IND	indicative

978–0–19–920647–6 00-Spencer-Prelims OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) xiii of xxviii May 16, 2008 19:1

INDEF	Indefinite
INESS	inessive
INF	infinitive
INGR	ingressive
INS	instrument(al)
INT	interrogative
INTENT	intentive
INTR	intransitive
IPFV	imperfective
IPS	impersonal
IRR	irrealis
ITER	iterative
KIN	kinship term
KP	Case Phrase
L	ligature (marker)
LAT	lative
LD	Locative-Directive marker
LFG	Lexical Functional Grammar
LOC	locative
LOCUT	locuter (1,2 p.)
M	masculine
MABL	'modal' ablative
MAT	Material (function/marker)
MEA	Means (function/marker)
MNOM	marked nominative
MOBL	modal oblique
MP	Minimalist Program
MPROP	'modal' proprietive
MR	macrorole
N	neuter
N	nominal/noun
N.PRED	nominal predicate
NARR	narrative
NEG	negative
NFUT	non-future
NHUM	non-human
NM	noun class marker
NMR	non-macrorole
NOM	nominative
NP	noun phrase
NPRS	Non-present
NPST	Non-past

NUC	nucleus
O	transitive object
OBJ	Objec(tive) case or marker
OBL	oblique (case)
ORI	orientative
OS	object–subject
OT	Optimality Theory
P	transitive object-patient
PAIP	Primary Actant Immunity Principle
PAM	Person Agreement Marker
PART	partitive
PASS	passive
PAST	past tense
PAT	patient(ive) case
PCL	particle
PEE	possessee
PEG	Pegative (case)
PERFREL	perfective relative
PERS	person marker
PF	P(atient)-Focus (verb form)
PFV	perfective
PHR:TERM	phrase terminal marker
PL; pl	plural
PN	proper noun
POL	Polite
POR	possessor
POSS	possessive (marker/pronoun)
POST	localization behind the landmark
POT	potential (mood)
PP	prepositional phrase
PRED	predicate/ predicative
PREP	preposition
PRES	present tense/ 'presentative' case
PRET	preterit
PREV	preverb
PRF	perfect
PRIV	privative
PROG	progressive
PROL	prolative
PROP	proper
PROPR	proprietive
PSA	privileged syntactic argument

PTCP	participle
PURP	purposive
PV	P(atient)-voice (?)
Q	question particle
R	recipient
RA	ra (accusative/oblique marker)
RDP	reduplication
REC	recent past
RECIP	reciprocal
REF	referential
REFL	reflexive
REL	relative pronoun, marker
RG	Relational Grammar
RM	remote (tense marker)
RN	relational noun
RRG	Role and Reference Grammar
S	(intransitive) Subject (marker)
S	sentence
SAP	speech-act participant
SBJV	subjunctive
SF	S-Focus
SG; sg	singular
SH	subject honorific
SMP	subject-marking particle
SO	Subject-Object
SPC	Specific(ity) marker
SPON	spontaneous voice
SRC	source (function)
SRESS	superessive
SS	same-subject marker
STAT	Stative
SUB	localization under the landmark
SUBJ	subject
T	ditransitive object-theme
TAM	tense-aspect-mood formant
TC	topic-contrast particle
TEMP	temporal
TLNK	topic linker
TNS	tense
TOP	topic
TP	Tense Phrase
TR	transitive

TRNSF	transformative
U	undergoer
UG	Universal Grammar
USIT	usitative
V	verb
VDAT	verbal dative
VEN	venitive
VOC	vocative
VP	verb phrase
WALS	Haspelmath et al. *The World Atlas of Language Structures*

Acknowledgements

A number of people have helped us in implementing this project. The original idea for the Handbook arose in the autumn of 2005, when Andrew Spencer was a guest of the PIONIER project, supervised in Nijmegen by Helen de Hoop. We would like to thank Helen for her unfailing support: she has been a major driving force behind the project from the first stages. Many of the contributors to the Handbook have been associated with the PIONIER project, either directly as project members (e.g. Andrej Malchukov), or indirectly, as guests and/or conference participants. Andrej Malchukov gratefully acknowledges the financial support of the Netherlands Organisation for Scientific Research (NWO) for the PIONIER project 'Case cross-linguistically' (grant 220-70-003). Andrej Malchukov also thanks Bernard Comrie and Martin Haspelmath, who provided ideal conditions for completion of the project in its final stages when he moved to Max Planck Institute for Evolutionary Anthropology, as well as acknowledging the financial support of DFG for the project on the typology of ditransitive constructions. Thank you to John Davey at OUP for his encouragement and patience, as well as Chloe Plummer and Julia Steer, and thanks also to Malcolm Todd for his quietly efficient copy-editing. Finally, many thanks to Birgit Jaenen for her last minute help in compiling the consolidated bibliography and to XXXXXX XXXXXXX for compiling the index.

We would like to thank our authors for their collegiality and patience and for their willingness to revise their articles so as to comply with our word limits (an important practical necessity in a Handbook with such a large number of chapters). We are grateful both to internal and external reviewers of the volume. Special thanks go to those who provided us with advice on editorial matters and helped us to improve the structure in various ways, including John Anderson, Barry Blake, Jim Blevins, Miriam Butt, Denis Creissels, Martin Haspelmath, Helen de Hoop, Edith Moravcsik, and Anna Siewierska.

Finally, a word of very special thanks to Larissa and to Marina.

Andrew Spencer
Andrej Malchukov

April 2008

The Authors

Mengistu Amberber is a Senior Lecturer in Linguistics at the University of New South Wales (Sydney, Australia). His research interests include lexical semantics, the semantics-syntax interface, morphosyntactic theory, and linguistic typology. He is the co-editor of *Language Universals and Variation* (Praeger, 2002), *Competition and Variation in Natural Languages: The Case for Case* (Elsevier, 2005), and editor of *The Language of Memory in a Crosslinguistic Perspective* (Benjamins, 2007).

John M. Anderson is interested in the nature of syntactic and phonological categorization, and in their relationship to the structures that relate the categories. His books include *Modern Grammars of Case* and *The Grammar of Names* (OUP, 2006 and 2007).

Peter M. Arkadiev is a Junior Research Fellow at the Department of Typology and Comparative Linguistics, Russian Academy of Sciences. His scholarly interests are linguistic typology, grammatical categories, and the semantics-syntax interface (including such topics as morphological case, event structure, and aspect). His 2006 PhD dissertation (written in Russian) was entitled 'The typology of two-term case systems'.

Markus Bader is a lecturer in linguistics at the University of Konstanz. He received his PhD at the University of Stuttgart in 1995 and has worked at the Universities of Jena, Konstanz, and Massachusetts at Amherst. His research interests include human language comprehension, in particular parsing and interpretation, and syntactic theory.

Matthew Baerman is a research fellow in the University of Surrey Morphology Group. His research focuses on inflectional morphology. He has published in *Language*, *Journal of Linguistics*, *Yearbook of Morphology*, *Linguistic Typology*, *Natural Language and Linguistic Theory*, and *Russian Linguistics*. He is co-author of *The Syntax-Morphology Interface: A Study of Syncretism* (CUP, 2005), and co-editor of *Deponency and Morphological Mismatches* (OUP, 2007).

Dik Bakker is Associate Professor of General Linguistics at the University of Amsterdam. His interests include language contact and change, functional approaches to grammar, and, in cooperation with Anna Siewierska, language typology. He has developed computational tools for the exploration of typological databases and bilingual text corpora and language sampling.

Jóhanna Barðdal is a Senior Research Fellow in General Linguistics at the University of Bergen. She works on syntax and semantics in Germanic languages, in a cognitive-functional framework. She has published articles in *Language*, *Cognitive Linguistics*, and J*ournal of Linguistics*, and is currently running a comparative project on non-canonical case marking in the early and archaic Indo-European languages.

Balthasar Bickel is Professor of General Linguistics at the University of Leipzig. His research interests centre on the worldwide distribution of linguistic diversity, its historical causes, and its relationships to social and cognitive processes. Since the early 1990s he has been engaged in fieldwork on typological outlier languages in the Himalayas. Together with Johanna Nichols, he is co-director of the AUTOTYP research programme for typological databasing. More information can be obtained at www.uni-leipzig.de/∼bickel.

Barry Blake retired in 2003, after four decades, from the position of Foundation Professor of Linguistics at La Trobe University. His publications include *Relational Grammar* (Routledge, 1990), *Case* (CUP, 1994, 2001), *Playing with Words* (Equinox, 2007) and *All about Language* (OUP, 2008) as well as several works on Australian Aboriginal languages.

James P. Blevins is Assistant Director of Research at the Research Centre for English and Applied Linguistics in the University of Cambridge. He received his PhD in Linguistics from the University of Massachusetts, Amherst in 1990 and has published on a range of grammatical topics. His research deals mainly with the analysis of morphological systems and syntactic constructions, with an emphasis on paradigm structure and discontinuous dependencies. Areal interests include Germanic, Balto-Finnic, Balto-Slavic, and Kartvelian.

Jonathan David Bobaljik is an Associate Professor of Linguistics at the University of Connecticut, Storrs. His major research interests lie in the areas of morphology and syntax, and their interface. He also has interests in language documentation and preservation, especially of the indigenous languages of the arctic and sub-arctic where he has conducted fieldwork.

Miriam Butt is a professor of theoretical and computational linguistics at the University of Konstanz. She works primarily on Urdu (syntax, morphology, semantics, some phonology) and is currently concentrating on the history and distribution of case. She is also working on building a computational grammar (parser/generator) and finite-state morphology for Urdu.

Denis Creissels is Professor of Linguistics at the University of Lyon. His main interests are descriptive linguistics and syntactic typology. His recent publications include 'Suffixes casuels et postpositions en hongrois' (*Bulletin de la Société de Linguistique de Paris*, 2000), and 'Encoding the distinction between location, source

and destination: a typological study' (*Space in Languages*, edited by Hickman and Robert, 2006). He is currently writing a grammar of the Nakh-Daghestanian language Akhvakh.

Michael Daniel is a researcher in linguistics at the Institute for World Culture Linguistic Ecology Department, Moscow State University. His interests include Armenian corpus linguistics and the typology of nominal morphology.

Alan Dench's principal expertise lies in the primary documentation and grammatical description of Australian Aboriginal languages, especially those of Western Australia. He has written grammatical descriptions of three languages of the Pilbara – Panyjima, Martuthunira, and Yingkarta – and is working on a description of Nyamal. He has also made contributions to the historical and comparative analysis of Australian languages, to studies of language contact, and has written in the general area of ethno-linguistics.

Mark Donohue researches the languages and linguistic prehistory of Indonesia and New Guinea. He has published a descriptive grammar of Tukang Besi and is currently working on a formal semantic account of its major syntactic phenomena. In addition to work in synchronic and diachronic Austronesian linguistics, he works on Papuan languages and the interaction between Austronesian, Australian, Papuan, and mainland. He is co-editor of *The Typology of Semantic Alignment* (OUP, 2008).

Sonja Eisenbeiss is a lecturer in psycholinguistics at the University of Essex. She has published on morphological processing, impaired and unimpaired first language acquisition, and the relationship between language and cognition. She has conducted experimental studies with children and adults and has developed games for the elicitation of various grammatical markers and constructions. Currently, she is analysing her extensive video-based corpus of German child language, funded by the Max-Planck-Society, with respect to children's morphosyntactic development.

N. J. Enfield is at the Max Planck Institute for Psycholinguistics in Nijmegen. He conducts fieldwork-based research in mainland Southeast Asia on micro and macro dimensions of language and other modes of signifying. Books (edited and authored) include *Ethnosyntax* (OUP, 2002), *Linguistic Epidemiology* (Routledge, 2003), *Roots of Human Sociality* (Berg, 2006, co-edited with S. C. Levinson), *Person Reference in Interaction* (CUP, 2007, co-edited with T. Stivers), and A Grammar of Lao (Mouton, 2007).

Dmitry Ganenkov is a researcher in the Department of Caucasian Languages, Institute of Linguistics, Moscow. His published work includes 'Experiencer coding in Nakh-Daghestanian' in *Case, Valency and Transitivity* edited by Leonid Kulikov, Andrej Malchukov, and Peter de Swart (Benjamins, 2006).

Martin Haspelmath is a senior staff member at the Max-Planck-Institut für evolutionäre Anthropologie in Leipzig. He received degrees at the Universität zu Köln, University at Buffalo, and the Freie Universität Berlin. He taught at the FU Berlin, the Universität Bamberg, the Università di Pavia, the Universität Leipzig, and at various summer schools. His research interests are primarily in the area of broadly comparative and diachronic morphosyntax. He is one of the editors of *The World Atlas of Language Structures* (OUP, 2005).

Bernd Heine is Emeritus Professor at the Institute of African Studies (Institut für Afrikanistik), University of Cologne. His 33 books include *Cognitive Foundations of Grammar* (OUP USA, 1997); with Derek Nurse, *African Languages: An Introduction* (CUP, 2000), *A Linguistic Geography of Africa* (CUP, 2007); 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 Genesis of Grammar* (OUP, 2007); and, edited with Heiko Narrog, *The Oxford Handbook of Linguistic Analysis* (in prep.).

Helen de Hoop is an Associate Professor in Linguistics at the Radboud University Nijmegen. Her theoretical and applied work in syntax and semantics focuses on case and interpretation. Her PhD dissertation *Case Configuration and Noun Phrase Interpretation* (Groningen,1992) was published in 1996 in the Garland series Outstanding Dissertations in Linguistics. In 2002 she was awarded a PIONIER-grant by the Netherlands Organisation for Scientific Research (NWO) for the research programme 'Case Cross-linguistically' to investigate the relation between morphological case and meaning.

Oliver Iggesen has been a research fellow at the Research Centre for Linguistic Typology since 2005. His current project is a descriptive grammar of Chacobo, a Panoan language spoken in the Amazonian lowlands of Bolivia. He received his PhD in 2005 from the University of Bremen with a dissertation on case-asymmetry, based on a large-scale cross-linguistic study involving a sample of 260 languages. His main interests are language typology, language contact, grammaticalization, and historical linguistics.

Lars Johanson is Professor of Turcology at the Oriental Institute of the University of Mainz, Germany. He has published extensively on synchronic and diachronic linguistics, especially in the domains of aspect-mood-tense, language contact, and language typology. Most of his publications focus on the Turkic language family. He is the Editor-in-Chief of the journal *Turkic Languages* and the book series *Turcologica*.

Seppo Kittilä is a university lecturer in general linguistics at The University of Helsinki. His main interests in linguistics include linguistic typology, syntax, argument marking, and transitivity. Recently he has been working on ditransitive

constructions, beneficiaries, and animacy effects on grammar. He has published articles on these topics, for example in *Linguistic Typology*, *Studies in Language*, and *Linguistics*.

Christa König is Professor of African Languages at the University of Frankfurt. Her research interests include verbal aspect and case systems. She has carried out extensive field research on Maa (Kenya, Tanzania), Ik (Northeastern Uganda), and !Xun (Namibia), and is presently engaged on a grammar of the Khoisan language !Xun. Her main book publications are *Aspekt im Maa* (AMO, 3, Cologne, 1994), *Kasus im Ik* (Nilo-Saharan Studies, 16, Köppe, 2002), and *Case in Africa* (OUP, 2008). She is the editor of *Afrikanistische Monographien*.

Leonid Kulikov is a post-doctoral researcher at the Leiden University Centre for Linguistics, currently working on a diachronic syntactic dictionary of the Old Indo-Aryan verb, as well as on a grammar of early Vedic (the language of the Rgveda). His research interests include synchronic and diachronic typology, in particular, typology of voices and valency-changing categories, as well as Sanskrit (Vedic) verb and syntax. He is co-editor of *Tense-aspect, Transitivity and Causativity* (1999) and *Case, Valency and Transitivity* (2006).

Monique Lamers is a post-doctoral researcher in linguistics, Radboud University Nijmegen. She studied Speech and Language Pathology in Nijmegen and Neurolinguistics in Brussels (Vrije Universitiet, Belgium). She carried out her PhD research in Groningen where she defended her thesis 'Sentence processing: using syntactic, semantic and thematic information' in 2001. From 2000 to 2003 she was a postdoctoral researcher in Magdeburg, Germany. She is currently a co-director of the 'Animacy' project at Radboud University in which her research focuses on the role of animacy in language comprehension and production, including differences between the intact and impaired language system.

Yury A. Lander is a research fellow in Asian and African languages, Institute of Oriental Studies RAS, Moscow. He specializes in North Caucasian languages, but also works on the Austronesian family. His interests include possessive, quantificational, and relative constructions, and the issues related to the description of polysynthetic languages. He is a co-editor (with Ji-yung Kim and Barbara H. Partee) of *Possessives and Beyond: Semantics and Syntax* (2004).

Silvia Luraghi teaches general and Indo-European linguistics at the University of Pavia. She has an MA in Classics from the University of Turin and a PhD in linguistics from the University of Pavia. Her research interests include historical and Indo-European linguistics, with a focus on Latin, ancient Greek, and Hittite, linguistic typology and cognitive linguistics. She has worked at the University of Chicago and at the University of Rome 3.

Andrej Malchukov is a senior researcher at the Institute of Linguistic Studies at the Russian Academy of Sciences, St Petersburg, and is currently affiliated to

Max Planck Institute of Evolutionary Anthropology in Leipzig. He specializes in Tungusic languages and has written a habilitation thesis on aspects of the morphosyntax of Even. His main interest is in morphosyntactic and semantic typology. Recent publications include *Nominalization/Verbalization: Constraining Typology of Transcategorial Operations* (2004). He is currently working on the typology of ditransitive constructions with Bernard Comrie and Martin Haspelmath.

Joan Maling is Professor of Linguistics at Brandeis University in Massachusetts, USA. Her doctoral dissertation in Linguistics at MIT reformulated al-Xaliil's metrical system of classical Arabic poetry in modern generative terms. She has published on many aspects of the syntax of modern Icelandic, especially case, word order, passive, preposition-stranding, and long distance reflexives, and on case alternations in Icelandic, Finnish, Korean, and German. She is currently investigating the development of a new impersonal passive construction in modern Icelandic. She was a founding editor of the journal *Natural Language and Linguistic Theory* and served as its managing editor for twenty-five years.

Elena Maslova is the author of *A Grammar of Kolyma Yukaghir* (2003) and *Tundra Yukaghir* (2003) and the editor of a linguistically analysed collection of Yukaghir texts (2001). She has worked on the typology of information-structure phenomena, reciprocal constructions, and case systems. Her current research interests focus on the methodological foundations of empirical typology, especially on analysis and interpretation of cross-linguistic and intra-genetic distributions of linguistic phenomena.

Alissa Melinger is a lecturer in psychology at theUniversity of Dundee. She previously worked at the MPI for Psycholinguistics in Nijmegen and in Computational Psycholinguistics at Saarland University. She is primarily interested in language production, focusing on the interface between lexical and syntactic processes during sentence production.

Edith Moravcsik teaches linguistics at the University of Wisconsin-Milwaukee (USA). Her main interests are syntax, language typology, cognitive linguistics, and Hungarian grammar. Her books include *An Introduction to Syntax*: *Fundamentals of Syntactic Analysis* and *An Introduction to Syntactic Theory* (both Continuum, 2006).

Åshild Næss is a postdoctoral researcher at the University of Oslo. Her main research interests are language typology and descriptive linguistics. She has worked on the Äiwoo and Vaeakau-Taumako languages of the Reef Islands, Solomon Islands, and has studied transitivity and case-marking from a typological perspective. Her published work includes *Prototypical Transitivity* (2007), descriptive and typological papers in scholarly books and journals, and collections of narrative texts in Äiwoo and Vaeakau-Taumako.

Bhuvana Narasimhan is Assistant Professor of Linguistics at the University of Colorado, Boulder. She conducts research on verb meaning, argument structure,

and information structure from a cross-linguistic and developmental perspective. Current research is based on videotaped corpus and elicited production data from children and adults speaking a variety of languages including Hindi, Tamil, Dutch, and German.

Heiko Narrog is an associate professor at Tohoku University, Japan. He has published on diachronic and semantic issues in verbal categories, especially modality, and on issues in language typology. His recent work includes papers in *Cognitive Linguistics*, *Diachronica*, *Language Sciences*, and *Sprachtypologie und Unversalienforschung*. He is co-editor with Heiko Narrog of *The Oxford Handbook of Linguistic Analysis* (in prep.).

Ad Neeleman is Professor of Linguistics at UCL. He has worked on a range of topics, including case theory, information structure, the encoding of grammatical dependencies, and the syntax-morphology interface. He is a co-author of *Flexible Syntax* (1999; with Fred Weerman) and *Beyond Morphology* (OUP, 2004; with Peter Ackema). His published work includes articles in *Linguistic Inquiry*, *Natural Language and Linguistic Theory*, *Linguistic Review*, and *The Yearbook of Morphology*.

Johanna Nichols is Professor of Slavic linguistics at the University of California, Berkeley. Her research interests centre on the worldwide distribution of linguistic diversity and complexity and their implications for prehistoric migrations and language spreads. She works on Slavic and other language families of the western Eurasian steppe periphery and on the Ingush and Chechen languages of the Caucasus. Together with Balthasar Bickel, she is co-director of the AUTOTYP research programme for typological databasing (www.uni-leipzig.de/∼autotyp).

Akio Ogawa is Professor of German Linguistics at the Kwansei Gakuin University, Nishinomiya. His main interests are language typology with special reference to contrastive studies on Japanese and German. He works on case-marking systems, especially dative constructions, as well as voice mechanism including passive, reflexive, and causative constructions, and impersonal constructions in their (dis)continuum to personal ones.

Enrique L. Palancar is Professor of Philology and Linguistics at the Universidad Autónoma de Querétaro. His major research interests include language typology, historical linguistics, syntax, and language description. He has a special interest in indigenous languages of Mesoamerica as well as old Indo-European languages, and is currently involved in the documentation of Otomi languages. Publications include *The Origin of Agent Markers* (2002) and journal articles on Otomi grammar and other topics.

Sandra Pappert is a lecturer in linguistics at the University of Leipzig. Her research focuses on sentence processing in German, covering comprehension and production as well as cross-modal processes.

Thomas Pechmann is Professor of Psycholinguistics at the University of Leipzig. He previously worked at the Max Planck Institute for Psycholinguistics in Nijmegen and at the Universities of Saarbruecken and Marburg. He is mainly interested in language production, in particular the activation of grammatical features during lexical access, and in the cognitive psychology of music.

Beatrice Primus is Professor of Linguistics at the University of Cologne and President of the Scientific Board of the Institute of German Language, Mannheim. She is co-editor of the monograph series *Linguistische Arbeiten* (Niemeyer). Her major areas of research are semantic roles, case marking, word order, and writing systems.

Esther Ruigendijk is an assistant professor of Dutch linguistics at the University of Oldenburg. Her research interests include language impairment, language processing, and language acquisition with a cross-linguistic approach. She is currently focused on the comprehension and production of pronouns and determiners. Her dissertion, *Case assignment in agrammatism: a cross-linguistic study*, was awarded a PhD by the University of Groningen in 2002.

Masayoshi Shibatani is Deedee McMurtry Professor of Humanities at Rice University and Emeritus Professor of Linguistics at Kolb University. He has conducted research on voice systems for many years and is currently investigating Austronesian Voice Systems in Eastern Indonesian. His wide interests in language include language typology, language universals, syntax and semantics, Japanese linguistics, and English grammar.

Anna Siewierska has been Professor of Linguistics and Human Communication in the Department of Linguistics and English Language at Lancaster University since 1994. Her major research interests are language typology, the comparison of different theoretical frameworks, diachronic change, discourse pragmatics, and most recently the morphosyntax of English dialects. She is working on impersonal constructions and (with Willem Hollmann) developing a project on morphosyntactic saliency in the context of their investigation of Lancashire dialect.

Andrej Sobolov is Professor of Slavic Linguistics at the Phillipps University, Marburg, where he was first appointed in 1997 as a lecturer in Slavic philology and Balkan linguistics. He was worked extensively on languages such as Old Church Slavonic, Russian, Serbian, Bulgarian, Macedonian, Czech, Polish, Albanian, and Greek.

Andrew Spencer is Professor of Linguistics at the University of Essex. He is the editor, with Arnold Zwicky, of *The Handbook of Morphology* (Blackwell, 1998) and the author of *Phonology*: *Description and Analysis* (Blackwell, 1996) and *Morphological Theory* (Blackwell, 1991).

Don Stilo is a Visiting Scholar in Linguistics at the Max Planck Institute for Evolutionary Anthropology, Leipzig, where he was a full-time Scholar 2001–6. He was also a Fellow at the Swedish Collegium for Advanced Study in Uppsala in 2005–6. He has a BSc from Georgetown University, and an MA and PhD from University of Michigan. He taught Persian language and Iranian linguistics for twenty years.

Thomas Stolz is chair of linguistics at the University of Bremen. PhD (Bochum, 1985). His research activities include morphology, language change, language contact, areal linguistics, typology, and universals. He has worked and published on Icelandic, Welsh, Latvian, Classical Aztec, Chamorro, and Maltese. He is editor-in-chief of the journal *Sprachtypologie und Universalienforschung* (STUF) and of the series *Diversitas Linguarum*.

Cornelia Stroh works at the linguistics department of the University of Bremen. She has collaborated in a variety of typological projects and is currently editorial assistant for the journal *Sprachtypologie und Universalienforschung* (STUF) and of the series *Diversitas Linguarum*. She has published on issues of language awareness on the German-French border (PhD, Bremen), and various topics related to the comitative and other categories.

Peter de Swart has written a doctoral dissertation on cross-linguistic variation in direct object marking. His main research interests include animacy, case alternations, and transitivity both from a typological and a formal linguistic point of view. Currently he teaches and is working on the influence of animacy on morphosyntactic patterns in the languages of the world within the NOW-financed research project 'Animacy' at the Radboud University Nijmegen, the Netherlands.

Aina Urdze has collaborated in several typological projects (on comitatives and alienability) at the University of Bremen. She has a degree in Finno-Ugric Philology from the University of Hamburg. She is currently finishing her PhD thesis on a class of verbs in her native Latvian. She has co-authored a variety of publications on dependent-marking in Latvian, comitatives and abessives, etc.

Robert D. Van Valin, Jr is Humboldt Professor of Linguistics at the Heinrich Heine University of Düsseldorf, before which he was for many years Professor and Chair of Linguistics at the University at Buffalo. He is the principal architect of Role and Reference Grammar. His research focuses on syntactic theory, language acquisition, and American Indian languages. His books include *Syntax: Structure, Meaning, and Function* (CUP, 1998); *Advances in Role and Reference Grammar* (Benjamins, 1993); and *An Introduction to Syntax* (CUP, 2001). He is the general editor of *Oxford Surveys of Syntax and Morphology*.

Maria D. Voeikova is the Head of the Department of Grammar Theory at the Institute of Linguistic Studies, Russian Academy of Sciences, St Petersburg. She teaches Russian morphology and psycholinguistics courses at the Faculty of Philology and

Fine Arts in St Petersburg State University. Her research interests and publications concern the functional description of Russian (Be-sentences, expression of quality, performatives, comparative studies) and the acquisition of Russian (quality and quantity, early vocabulary, diminutives, pre- and protomorphology, typological differences).

Fred Weerman is Professor of Dutch Linguistics at the University of Amsterdam. In his research he combines theoretical linguistics with work on language change and language acquisition. His books have been published by Foris, Kluwer, and OUP. His articles include work published in *Folia Linguistica Historica*, *Language Acquisition*, *Lingua*, *Linguistics*, *Natural Language and Linguistic Theory*, *Nederlandse Taalkunde*, *Morphology and Second Language Research*.

Søren Wichmann is a postdoctoral fellow at Max Planck Institute for Evolutionary Anthropology and Leiden University. He has published on aspects of historical linguistics (grammaticalization, borrowing patterns, ancient scripts, quantitative methods, computational modelling) and has contributed to documenting Tlapanec and Texistepec Popoluca. Recent papers have appeared in journals such as *Linguistic Typology*, *Diachronica*, *Annual Review of Anthropology*, and *Studies in Language*. He is co-editor of *The Typology of Semantic Alignment* (OUP, 2008).

Anna Wierzbicka is Professor of Linguistics at the Australian National University. Her work spans a number of disciplines, including anthropology, psychology, cognitive science, philosophy, and religious studies as well as linguistics. She is the author of numerous books, most recently, *English*: *Meaning and Culture* (2006). Earlier publications include *Cross-Cultural Pragmatics* (de Gruyter,1991, 2nd edition 2003), *Semantics*: *Primes and Universals* (OUP, 1996), *Understanding Cultures through their Keywords* (OUP, 1997), and *Emotions across Languages and Cultures*: *Diversity and Universals* (CUP,1999).

Susanne Wurmbrand is an Assistant Professor of Linguistics at the University of Connecticut, Storrs. Her major theoretical research interests are syntax and the syntax-semantics interface and recent publications include *Infinitives: Restructuring and Clause Structure* (Mouton de Gruyter, 2003). She is interested in all aspects of Germanic linguistics, and is the executive editor-in-chief of the *Journal of Comparative Germanic Linguistics*.

Joost Zwarts worked as a researcher in the PIONIER project 'Case Cross-linguistically' at Radboud University Nijmegen. He now teaches at the department of Linguistics at Utrecht University. His main interests are in (formal and lexical) semantics and syntax, especially in the behaviour of prepositions and articles. He has done fieldwork in Kenya and published *The Phonology of Endo* (Lincom, 2004).

INTRODUCTION

ANDREJ MALCHUKOV
ANDREW SPENCER

The notion of 'case' has played an important role in thinking about grammar since the days of Pāṇini and Aristotle. Nonetheless, the concept of case and its relation to grammatical relations, meaning, and morphological form remains elusive and controversial. In modern times whole grammatical frameworks have been developed taking some concept of 'case' as a central core, and at the same time, other types of theoretical approach have taken on the challenge of assimilating and explicating the notion, whilst typological studies have been exploring the variety of the phenomena commonly considered under the rubric of 'case'. At the same time some researchers have sought to bridge the gap between typological or descriptive studies and theoretical studies (the work of authors such as Comrie and Mel'čuk comes to mind; see e.g. Comrie 1986; Melčuk 1986). More recently, monographs and extensive research projects have been devoted to the problem, such as the books by Blake (whose first edition appeared in 1994) and more recently Butt (2006). At the same time two major European research projects have been devoted to case in recent years: 'Case and thematic relations' at the Leuven University (see e.g. Davidse and Lamiroy 2002), and the PIONIER project 'Case cross-linguistically' (supervised by Helen de Hoop) at the Radboud university Nijmegen. The problems and challenges of case systems in the languages of the world continue, therefore, to exert a fascination on linguists and other specialists from a variety of backgrounds, and this is reflected in the rise in the number of monographs and general introductions to the subject.

However, this rise in interest in case brings with it difficulties for anyone wanting to keep abreast of latest findings and ideas, since there is currently no single

978–0–19–920647–6 01-Spencer-Intro OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 2 of 10 May 16, 2008 18:57

extended source of information about case that scholars can turn to. It was this consideration that gave rise to the idea for this Handbook, while Andrew Spencer was a guest of the PIONIER project in the autumn of 2005. We felt that there was a need for a single volume which would bring together in synopsis form recent work on the problems of case, focusing on as many relevant aspects as possible. We have invited scholars from a variety of different backgrounds and theoretical persuasions to summarize the way that the notion of case figures in current grammatical theory, and how it relates to other aspects of morphology, syntax, and semantics. The Handbook also contains articles on the types of case systems that languages exhibit and the way that case paradigms are structured, the way case systems develop and decay over time, and the kinds of functions and meanings that are expressed by case systems. In addition, the Handbook offers the reader a variety of articles relating the notion of case to other grammatical phenomena such as transitivity, the alignment of grammatical relations, and so on. In recent years psycholinguistic studies have focused on the way that case systems are acquired by language learners, or lost in language disorders, and the way that case systems are processed on-line by adult language users, and so these topics are also included. We hope, therefore, to have provided a convenient starting point for any student or researcher who wishes to gain an entré into the world of case as well as providing detailed summaries of specific phenomena which are otherwise only discussed in relatively inaccessible places.

The phenomenon of grammatical case has been a central feature of the Western grammatical tradition for some two millennia because of the importance of case in Latin and Greek. Equally, the Paninian tradition of Sanskrit grammar writing has laid great stress on the notion of kāraka, a notion which is related to, but not identical to the Graeco-Latin conception of case. In a sense there are two notions associated with 'case' which, whilst closely related, sometimes have to be distinguished. One is the formal notion of case as an inflected form of a nominal word and the other is the 'semantic' notion of case as a function of a nominal phrase in another phrase or in a clause. One can perfectly well talk about 'case functions' at an abstract level without necessarily relating this to any particular kind of morphology. As a result, we often see studies of case which discuss the grammatical behaviour of words or phrases that are marked with prepositions or postpositions rather than purely inflectional case markers, or indeed phrases that are not marked at all, but are distinguished solely in terms of, say, word order. Some specialists in case would argue that this extension is not warranted, but we have decided to be as inclusive as possible in allowing authors to decide what is or isn't an example of 'a case'. In this way we hope we have not excluded phenomena which are important for current theoretical models or which might shed light on the typology of case.

The first set of substantive articles, Part I of the Handbook, is devoted to the way that case phenomena (broadly construed!) are treated in a variety of frameworks. It is appropriate that this part should start with an introductory chapter on the history

978–0–19–920647–6 01-Spencer-Intro OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 3 of 10 May 16, 2008 18:57

of research of case by Barry Blake, whose groundbreaking Cambridge book on case stimulated a new upsurge of interest in case, semantic roles, and grammatical relations. Blake's chapter is complemented by Miriam Butt's overview of modern approaches to case, summarizing some of the findings of her recent monograph on the theories of case (Butt 2006). John Anderson's chapter has two objectives. First, it introduces the localistic case theory that he has been developing over the past thirty-odd years (see also Anderson 2006), and second it provides an invaluable historical perspective on the evolution of case theories in the European tradition, thus bridging the gap between Blake's historical overview and Butt's chapter on current approaches.

In some instances the notion of case has played a substantial role in the development of the theoretical framework itself. This is particularly true of Anderson's localist case grammar, but case has also played a significant role in the development of Wierzbicka's Natural Semantic Metalanguage approach, and to some extent the related approaches within Cognitive Grammar summarized by Luraghi. Similarly, although optimality theoretic syntax has dealt with a whole host of issues, there is a significant subdomain which has applied Optimality Theory to a wide range of case-related problems, as discussed in de Hoop's chapter. The 'case-in-tiers' model discussed by Maling is exclusively a model of case marking, and hence is not really an example of a 'theoretical model', but its importance for the Handbook should be obvious.

The problems of case and related questions of alignment of nominal phrases with grammatical functions such as subject and object have played an important role in the development of both Lexical Functional Grammar, and Role and Reference Grammar, and the place of case in these models is presented by Butt and Van Valin respectively. A more abstract and in some ways more controversial notion of case ('Abstract Case') has played a significant role in the development of the various instantiations of the Principles and Parameters model of syntax, summarized in Bobaljik and Wurmbrand's article. The relationship between case functions and meaning is explored to varying extents in a whole host of articles in the Handbook, but the specific question of how case is viewed by formal semanticists is summarized by de Hoop and Zwarts.

Part II of the Handbook examines a variety of issues in the morphology of case. Spencer's article provides an overview of the kinds of complexities that case provides for morphological systems, such as linear ordering of case markers and cumulation with other categories. It argues for a distinction between formal or morphological case and syntactic case, on the basis of well-known mismatches between the two. In a related vein it raises the question of how we identify cases in the grammar of a language, appealing to 'Beard's Criterion', modified from Beard (1995). Some of these issues are discussed in more detail in the articles by Blevins and by Baerman. Blevins looks at the way that case systems distribute themselves in paradigms, arguing that we must look to the shape of the whole paradigm to

understand this organization. His arguments are complemented by Baerman who examines the specific question of syncretism. This refers to a situation in which a single word form has more than one meaning or function in the paradigm (so that a given word form might be both 'accusative case' and 'genitive case'). These two articles both show how case systems provide important material for those who wish to study paradigm systems generally. Moravcsik's article, by contrast, takes as its starting point the notion of a case marker (or 'case morpheme' if you will) and asks how these markers are distributed syntagmatically within word forms or within phrases generally. Her article and that of Spencer deal with related issues surrounding the linear ordering of case markers with respect to other elements, but they do so from complementary perspectives. Part II closes with a discussion of a specific distinction which is very common cross-linguistically. Case systems change over time, sometimes in drastic ways, but the effects on lexical nouns are often different from the effects on pronouns, which are extremely high frequency elements and therefore tend to be more conservative and less prone to change (witness English pronouns, which retain some vestiges of a now completely lost morphological case system). However, as Iggesen shows, pronouns, and other clearly delineated 'NP-types' exhibit a whole host of differences from normative paradigms in what he calls 'case-asymmetric' systems.

Part II is devoted to the syntactic functions and roles that cases fulfil. Primus' contribution sets the scene by providing an overview of the way that cases express grammatical relations (such as subject and object) and semantic roles (such as agent and patient). She surveys models based on an unanalysed list of semantic ('thematic', 'theta') roles, and models based on some kind of decomposition of semantic roles, particularly the influential proposals of Dowty ('proto-roles') and also the lexical decompositional approach, under which semantic roles are related to semantic primitive predicates such as CAUSE, BECOME, MOVE.

Like Bobaljik and Wurmbrand's chapter, the contribution by Neeleman and Weerman is coached within the generative tradition, with the difference that they focus on the role of morphological case rather than abstract case. Their chapter thus illustrates the way in which within a given theoretical framework both broad and narrow definitions of case may be appropriate, offering complementary perspectives. Neeleman and Weerman are principally concerned with word order effects and case, addressing the old, but still unclear, question of how the presence/absence of overt case marking might be related to free/fixed word order (in Dutch, German, Icelandic, and Japanese).

Siewierska and Bakker complement the discussion in Neeleman and Weerman's chapter with a broad typological survey of the relationship between case marking and the two other main encoding strategies, agreement morphology and word order. They examine the discriminating (or distinguishing) functions and the indexing (or 'characterizing') functions of each strategy (a distinction taken up in Malchukov and de Swart's chapter). The discriminating/distinguishing function

978–0–19–920647–6 01-Spencer-Intro OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 5 of 10 May 16, 2008 18:57

is that of distinguishing core grammatical relations such as subject from object. Indexing functions are those which relate to the semantics of the nominals marked by case, for example, where a case marker marks a subject provided it is animate or provided it has the semantic role of 'experiencer'. (These notions are further discussed in Primus' chapter.) Siewierska and Bakker point out that there is much variation in the use and combination of encoding strategies, but they note that there do exist statistical tendencies for certain types of encoding to correlate with certain types of function. For example, case marking tends to be correlated with basic word order type. (This is an issue that is also taken up in Malchukov and de Swart's chapter).

Bickel and Nichols focus on the way that case marking gets distributed across the principal grammatical roles ('alignment strategies'). On the basis of a meticulous examination of the available typological evidence the authors cast doubt on a number of widely-held beliefs about alignment. For instance, they fail to find support for the existence of so-called 'stative–active' alignment languages and they also question the role of the 'animacy hierarchy' and similar factors in determining whether a subject or object will be overtly marked. They conclude that what is really important is the complex set of lexical conditions that a language imposes on its case marking principles ('lexical valency sets').

Alignment strategies are closely related to the notion of transitivity, and this is a theme running through the last three contributions of Part III. Shibatani discusses the way case marking is reflected in valency alternations such as passives/antipassives (valency decreasing) and applicatives and external possessor constructions (valency increasing). He points out that passive and antipassive alternations are often effected by a combination of verb morphology and case marking, but that on occasions they can be expressed by verb morphology or by case marking alone (the latter possibility being illustrated, for instance, by the Samoan antipassive). Malchukov and de Swart examine the differential marking of subjects and objects, which they attribute to interaction of the two basic functions of case marking, distinguishing (discriminating) and differentiating ('indexing', 'characterizing'). They show how the interaction between the two functions can account for asymmetries between differential object marking and differential subject marking. Kittilä's chapter looks more specifically at the way case alternations interact with transitivity alternations. The chapter takes further a number of the issues addressed by Malchukov and de Swart, and together they can be thought of as explorations of the 'indexing' functions of case.

Ever since the pioneering work of Slobin on the acquisition of Russian and Turkish case specialists in first language acquisition have been interested in case, but more recently case has attracted the attention of researchers in other branches of psycholinguistics. Part IV provides an overview of the recent trends.

Eisenbeiss, Narasimhan, and Voeikova argue against strongly nativist approaches to case in child language, under which innate categories such as 'accusative' or

'absolutive' get mapped to innate conceptual semantic categories such as 'patient'. They contrast these approaches with usage-based models, which rely exclusively on general abilities to abstract regularities, and with models based on more articulated theories of case (specifically the Kiparsky/Wunderlich model summarized in Primus' chapter), in which case relations can be projected from the semantic relations which hold between participants on a word-by-word basis. They also provide an overview of the Natural Morphology model of W. U. Dressler and colleagues. The authors conclude that the mistakes children make are systematic, these patterns can't be accounted for purely in terms of semantics or of syntax.

Speech production research has tended to focus on English, which provides scant evidence for case marking. Melinger, Pechmann, and Pappert therefore complement the English data with work on German case marking. They review speech error data and on-line processing experiments, arguing that case assignment takes place at Garrett's 'functional' level, which is defined in terms of abstract representations of lexemes rather than the later 'positional' level, at which word forms are already specified. They also discuss the controversial issue of whether sentences are constructed incrementally or whether the overall structure of a clause is first determined by the choice of verb (even in verb-final clauses).

Bader and Lamers review evidence from reading experiments which investigate the role of case marking in Dutch, German, Korean, and Japanese for sentence comprehension, as well as extrapolating from English data. They focus on two possible functions for case marking: identifying grammatical roles of NPs and identifying clause boundaries (for instance, an unambiguously nominative pronoun in the middle of an English sentence usually marks the beginning of a finite clause). German readers prefer to interpret a NP as accusative rather than dative in neutralization contexts, raising the question of whether markedness effects can be discerned in processing. Finally, the authors summarize recent proposals for incorporating optimality theory into analyses of ERP studies ('event-related brain potential') on case processing in Dutch and German.

Lamers and Ruigendijk provide an overview of the way that case has figured in studies of acquired language impairment. Even Broca's aphasics, who are supposed to have a particular impairment in the processing of grammatical information, are sometimes able to make use of case morphology for disambiguation, though in general, processing of case suffers in these patients. Although they lament the comparative lack of studies on case in aphasic comprehension or production, they note that a number of Russian studies (replicated in studies of German and Hungarian) provide interesting insight into the impairment of case usage. Surprisingly, perhaps, the evidence seems to suggest that 'structural' case (i.e. nominative/accusative determined by overall clause structure) is less subject to error than 'lexical' case (which generally adds some component of meaning).

These three studies note that relatively little work has been devoted to case as such in normal or impaired language processing (as opposed to more general

phenomena such as grammatical function assignment, the computation of filler-gap dependencies, or the interpretation of reversible passive constructions). These authors all throw down interesting challenges to psycholinguists and aphasiologists with access to speakers of case-rich languages to investigate some of the phenomena discussed elsewhere in this Handbook, such as the distinguishing vs. indexing functions of case, abstract vs. lexical case, and more generally the issues of case markedness (see Malchukov and Spencer on the Case Hierarchy, Chapter 45), and a variety of other respects in which case might be reflected directly in sentence processing or language breakdown.

The chapters in Part V deal with related issues in the rise and fall of case systems and in the way they are distributed geographically. Kulikov illustrates the way the Indo-European case system was reconfigured in a variety of languages so as to reconstitute a largely lost case system (Indo-Aryan languages). He also outlines the mechanisms by which languages resist case change (e.g. in Armenian). Heine summarizes the grammaticalization paths found in the evolution of cases, for example, from spatial nouns to postpositions to locational cases. He also discusses the extensions of case meanings, often from more concrete to more grammatical meanings of functions. He also charts the way case markers acquire non-case functions, such as marking clause subordination, modal meanings, and so on. Barðdal and Kulikov trace the way cases are lost, through phonological attrition (as in Romance languages) but also through other pathways, as when complex patterns of argument structure marking get simplified in the history of Germanic (for example, the loss of non-nominative subject constructions).

Bickel and Nichols address the complex issue of areal distribution, adopting a sophisticated notion of 'linguistic area' that is, as far as possible, independent of language. They survey thirty-five case-related variables both morphological and syntactic using as source material various recently developed large databases ('Autotyp' and *World Atlas of Linguistic Structures* [WALS]; Haspelmath, Dryer, Gil, and Comrie 2005). Careful statistical analysis of the data reveals that 'aspects of position and fusion of case markers, their presence vs. absence, and their alignment are prone to areal spread while aspects of exponence, flexivity, syncretism, and phrasal behavior tend to resist spread'. Finally, Johanson continues the discussion of areal spreads looking at the way case oppositions or actual case markers are borrowed ('copied') from one language to another.

The last two parts of the Handbook present essentially descriptive/typological surveys. In Part VI we have summaries of the typical, and not-so-typical behaviour of individual cases (including the typologically unusual situation in which nominative is specially marked). This Part begins with Haspelmath's survey of case terminology (which also serves as a very useful introduction to the world's case systems). In addition to helping the reader navigate the bewildering variety of terms for the same thing and different uses of one and the same term, this chapter offers salutary warnings about the dangers of assuming that a case

label in one language denotes the same property as the same label in another language.

König discusses cases systems in which the accusative is morphologically and/or functionally unmarked compared with the nominative. This is largely, but not exclusively restricted to Africa, where it is the prevalent type of case system (especially where case is marked by tone). She further distinguishes between two types of marked nominative languages, depending on whether functional markedness is matched with the formal markedness in a case system or not.

Malchukov and Narrog show how the technique of 'semantic maps', common nowadays in functional-typological approaches, can throw light on case polysemy. This chapter serves also as an introduction to the spirit underlying several of the specific chapters on case in this Part. Several of the chapters of Part VI discuss interesting relationships between inflectional case proper and other coding strategies for case functions (for instance, Lander provides a useful survey of possessive marking generally, Narrog shows how adpositions and applicative verb forms fulfil instrumental functions, and so on). Creissels explores very rich systems of spatial cases, providing a wider typological context for the discussion of the highly elaborated systems of Daghestanian languages discussed in greater detail in Daniel and Ganenkov's chapter (the two chapters should be read in conjunction). He outlines the way such spatial cases develop, and how they develop into purely grammatical cases. He also touches on the special behaviour of nouns denoting places and other such restrictions on case systems. Several of the chapters take up issues addressed in the syntactic part, such as differential object marking (Kittilä and Malchukov's chapter on accusative, and Naess' chapter on the dative), and ergativity and differential subject marking (Palancar's chapter on the ergative). While all the chapters provide typological surveys, several (in particular that on the comitative by Stolz, Stroh, and Urdze, that on the instrumental by Narrog, and that on ergative by Palancar) report on the results of sample-based typological research uncovering areal patterns.

The vocative case is often excluded from typological or theoretical discussion of case systems, and Daniel and Spencer point out that in many languages the form of the vocative is distinctly unusual compared to other cases, and indeed may violate basic morphological or phonotactic principles elsewhere in the language. But they also point out that there are plenty of languages in which the vocative has been fully integrated into the case paradigm, and even triggers case agreement. Part VI concludes with Malchukov's round up of other less common properties of cases or case systems. He looks at cases with an unusual distribution, such as the 'case-stacking' (see also the chapters by Spencer and by Moravcsik), cases with unusual functions ('pragmatic cases'), and cases with both unusual functions and distributions, such as the modal cases and verbal cases of Kayardild. He argues, in particular, that cases with unusual properties often result from an incomplete grammaticalization cycle. Importantly, as shown throughout this Part, individual

978–0–19–920647–6 01-Spencer-Intro OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 9 of 10 May 16, 2008 18:57

cases display recurrent polysemy patterns with non-random functional overlap, which makes it possible to represent their functions in a single semantic network (see Malchukov and Narrog).

Part VII is largely devoted to descriptive summaries of specific types of case system, to give a flavour of the typological variation found. The scene for this is set by Malchukov and Spencer's overview of case typology, which raises a number of the issues addressed in individual contributions to this Part. Malchukov and Spencer look at case typology both at a 'macro-level', in the content of case paradigms across languages, and at a 'micro-level', in the relations between the set of morphological cases in a language and the set of syntactic cases. They also summarize the major functions (syntactic vs. semantic vs. pragmatic) that individual cases can subserve. At the 'macro-level', Blake has proposed that cases fall into a hierarchy which induces implicational generalizations: what is true of a case lower on the hierarchy, such as locative, should be true of cases higher on the hierarchy, such as dative or accusative. Malchukov and Spencer reassess the role of the case hierarchy in grammar. Drawing extensively on the contributions to the present volume, they show that the hierarchy assumes different guises in different linguistic and psycholinguistic domains. The next contributions contrast extremes: Daniel and Ganenkov survey the remarkably rich sets of cases found in many Daghestan languages, where it's common for languages to express notions such as 'from the surface of' or 'towards the front of' using case morphology. On the other hand, languages with just two cases are not particularly uncommon and Arkadiev examines the kinds of use to which such a minimal opposition can be put. Minimal two-case systems are often the result of attrition of larger systems (see Barðdal and Kulikov, ch. 30) and this is true of the depleted case systems of Iranian, discussed by Stilo, and of South Slavic languages studied by Sobolev. Sobolev traces the way case functions such as instrumental get marked by prepositions such as 'with' in Bulgarian and Macedonian. Stilo examines the way that the two-case systems common in Indo-Iranian languages have developed (or not) over time and examines the compensatory strategies in those languages that lost even the remnant two-term system. He also discusses the remarkably complex possibilities for variation in the alignment of case and grammatical functions in languages that show split ergativity as well as differential object marking. A particularly interesting additional strategy is noted for one of the Central Plateau Dialects, Gazi, in which a 'floating' subject agreement marker selectively cliticizes to the right edge of a direct object phrase. The case system of Ik described by König is one of the most unusual we have encountered, in that cases are marked on all parts of speech, though case has a very low discriminatory function with core participants (subjects/objects). Amberber's discussion of Amharic provides an in-depth illustration of the phenomenon of differential case marking discussed in various chapters in Part III, especially the chapter by Malchukov and de Swart. Dench's chapter on case in Nyamal is particularly focused on a phenomenon for which Australian languages are noted (though the

978–0–19–920647–6 01-Spencer-Intro OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 10 of 10 May 16, 2008 18:57

phenomenon is far from restricted to Australia), so-called 'case-stacking', in which an element with case marking which modifies a case-marked noun receives the case marker of that noun in addition to its own. Donohue's chapter discusses the case system of the Philippine language Tukang Besi, which picks out subject and object participants but in a manner which interacts in very complex ways with word order, agreement, and information structure, as a kind of microcosm of the Philippine case marking system. Nominal participants in Japanese also bear case markers that interact in complex ways with pragmatic/discourse functions, as Ogawa describes. Both Tukang Besi and Japanese provide examples of 'pragmatic cases' in the sense of Malchukov and Spencer, revealing a tight integration of semantic and discourse pragmatic information. The same is true of Yukaghir, discussed by Maslova, where (patient) case marking is sensitive to properties of information structure, and probably takes its origin in discourse (focus) markers. Wichmann's chapter describes the way that Tlapanec, an Oto-Manguean language of Mexico, uses verb affixes to code the sorts of functions normally coded by case markers on nominals in other languages (this language is also unusual in that its Azoyú dialect has a 'pegative' case function, indicating an actor in an event which also involves a Dative-like undergoer, in effect a unique distinguisher for the 'A2' function defined by Nichols and Bickel. Finally, Enfield shows how case-like functions are expressed in Lao, a language with very little morphology, and certainly no standard case morphology. His paper, like Wichmann's, highlights the similarities and differences between case proper and the alternative strategies of encoding grammatical relations (word order, head-marking) discussed by Siewierska and Bakker.

We hope that the Handbook will be useful as a reference source to general linguists, psycholinguists, computer scientists, and others with an interest in case systems, as well as being a source of inspiration for future research.

PART I

THEORETICAL APPROACHES TO CASE

CHAPTER 1

HISTORY OF THE RESEARCH ON CASE

BARRY J. BLAKE

1.1 The Greeks

Current western methods of analysing language follow a tradition that can be traced back to the Greeks. They developed criteria for determining word classes and terminology for describing word classes and grammatical categories. One of the main criteria was case, which for the Greeks was a system of word forms that signified relationships in sentences. The main word class to exhibit case forms was the noun, which for the Alexandrine Greeks included the adjective. Other word classes to show variation for case were pronouns, articles, and participles. In Ancient Greek there were five cases. The case forms of *anthrōpos* 'man' are shown in Table 1.1. There were three numbers – singular, dual, and plural – and there was fusion of case and number marking across the whole system, what Matthews calls cumulation (Mathews 1991).

The noun *anthrōpos* belongs to the second of three declensions. As can be seen from Table 1.1, all cases are distinguished in the singular, but there is some syncretism in the plural and even more in the dual. The cases in Table 1.1 are presented in the order used by the Greeks. Modern grammars usually adopt the order nominative, vocative, accusative, genitive, dative. This has a certain logic in

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 14 of 26 May 16, 2008 18:57

Table 1.1. Declension of *anthrōpos* 'man'

English	Greek	singular	dual	plural
nominative	*onomastikē*	anthrōpos	anthrōpō	anthrōpoi
genitive	*genikē*	anthrōpou	anthrōpoin	anthrōpōn
dative	*dotikē*	anthrōpō	anthrōpoin	anthrōpois
accusative	*aitiatikē*	anthrōpon	anthrōpō	anthrōpous
vocative	*klētikē*	anthrōpe	anthrōpō	anthrōpoi

that the core cases precede the others. Moreover, there is a lot of syncretism between nominative, vocative, and accusative, so putting these three together means putting like cases together.

Greek had a three-way gender distinction: masculine, feminine, and neuter. The majority of first declension nouns were feminine, almost all second declension nouns were masculine or neuter, and third declension nouns could be of any gender. Neuter nouns in the second declension declined like *anthrōpos,* but took *-on* for nominative-vocative-accusative singular and *-a* for nominative-vocative-accusative plural. Gender was manifest through the adjective. A large class of adjectives took first declension forms for feminine and second declension forms for masculine and neuter. Leaving aside a few adjectives that took masculine forms for feminine as well as masculine gender, adjective inflection was a true guide to gender and one could say there was fusion of gender with case and number. This fusion is perhaps one of the reasons the Greeks kept to a word-based model of description and never arrived at the concept of the morpheme.

The Greeks used the term *ptōsis* for 'case'. The word means 'falling' and originally referred to all the forms in which a root could appear be they inflected forms or derived ones. From the Stoic grammarians (third century BC) onwards the term was restricted to case forms. The nominative was referred to as the *ptōsis orthē* the 'upright' case or *ptōsis eutheia* 'straight' or 'direct' case, or later *ptōsis onomastikē* 'naming' or 'nominative' case. The other cases were called collectively *ptōseis plagiai* 'slanting' or 'oblique' cases. In chapter two of *Peri hermeneias* Aristotle declared that only the nominative is a noun, the other forms being cases of nouns. Later grammarians, including the Byzantines, often make the point that the nominative is not really a case since it presents just the referent, whereas the others incorporate some relation such as 'to' or 'from' the referent. This is not the whole truth. Where the accusative indicates extent in space or time as in *Emeine treis hēmeras* 'He remained three days', the case form incorporates a predicate meaning 'extent'. In poetry the accusative can express destination as in *Mnēstēras aphiketo* 'She came to the suitors' (prose would normally require the preposition *pros*); again the case form expresses a predicate, in this instance destination. But where the accusative expresses the direct object of a verb, it is merely a

marker distinguishing arguments of a predicate and its status is analogous to that of the nominative.

The oblique cases were named after salient functions. The genitive was the *ptōsis genikē*, where *genikē* refers to origin in the hereditary sense. The Greek genitive expressed 'from' as well as possession. It also marked the complement of various verbs including those expressing perception *phōnēs akouein* 'to hear a voice'. The accusative was the *ptōsis aitiatikē* where *aitiatikē* means 'causal'. This might seem strange, but what is meant is that the accusative indicates the patient of a caused action. The dative was the *ptōsis dotikē,* the 'giving' case, because it encoded the recipient of the verb *didōmi* 'to give'. In fact the dative had a variety of functions including a locative function, though mostly when governed by a preposition, and marking the complement of various verbs indicating aid, trust, pleasing, and the like.

The vocative was the *ptōsis klētikē* the 'calling' case, the case used in addressing someone. It was recognized that the vocative was a case purely because its inflection belonged to the case system. The vocative did not have a function like that of the other cases, which we could say was to mark the type of relation dependent nouns bore to their heads. Greek grammarians often omitted the vocative from their discussion of the Greek case system.

In any language dependent nouns will need to bear a greater variety of semantic relations than could be provided by a small case system, even allowing for some polysemy. In particular there will be a need to express local notions such as 'around', 'through', 'down', 'over', etc. In Greek the oblique cases had local meanings: the genitive expressed source of motion, the dative location, and the accusative expressed 'goal of motion'. However, the oblique cases were not much used for expressing these notions on their own. They were generally governed by prepositions which expressed a variety of local notions. Some prepositions governed one particular case. *Ek* (or *eks*) 'from', 'out of' governed the genitive: *ek tēs poleōs* 'from the city'; *en* 'in' governed the dative: *en Spartē* 'in Sparta'; and *eis* 'into' governed the accusative: *eis ton potamon* 'into the river'. Some prepositions governed more than one case. *Para* governing the genitive signalled 'from': *para theou* 'from God'; governing the dative it indicated 'near': *para Kurō* 'near Cyprus'; and governing the accusative it could mark 'to', 'along' or 'throughout': *para panta ton chronon* 'during the whole time'.

One Byzantine grammarian, Maximus Planudes (*c.* 1260–1310), was picked out by Hjelmslev (1935: 13–15) as having arrived at a localistic theory of case, that is, one in which various functions are interpreted in local terms, the agent, for instance, as a source, and the patient as a destination. In fact Planudes emphasized the local functions of the oblique cases, the fact that the genitive expressed 'from', the dative 'at', and the accusative 'to', primarily with reference to place and secondarily with respect to time. It is not clear that he took the further step of claiming the more abstract functions were derived from the local functions (Robins 1993).

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 16 of 26 May 16, 2008 18:57

1.2 The Romans

The Roman grammarians inherited the Greek tradition of grammatical description. They translated the Greek terminology into Latin so that *ptōsis* became *cāsus*, *onomastikē* became *nominatīvus*, *genikē* became *genitīvus*, *dotikē* became *datīvus*, and *klētikē* became *vocatīvus*. The *ptōsis aitiatikē* became the *accusatīvus*. In Greek *aitia* 'cause' and its derivatives had the extended senses of 'something/someone responsible' and 'the one accused'. Varro, a Roman grammarian of the first century BC, is thought to have been responsible for basing the Latin term on the 'accused' sense (Robins 1997: 44).

Latin had a case not found in Greek. It was given the name *ablatīvus*, the 'taking away' case. The ablative expressed the notion of 'from' and covered a range of meaning that could be related to 'from' such as expressing the agent of the passive. It also expressed the instrument with which an action was carried out. Quintilian (*c.* AD 35–95) suggested that a separate instrumental case be recognized. This suggestion was not taken up by later grammarians. Priscian, author of the *Institutiones grammaticae* (*c.* AD 500), for instance, rejected the idea.

The Romans essentially took over the Greek model of description, and found equivalent terms in Latin, most of which have become standard. One original contribution relevant to case can be found in Varro. He distinguished inflection from derivation, pointing out that inflection was regular in its distribution and always added a consistent meaning or combination of meanings, e.g. accusative plural, whereas derivation was irregular in its distribution and did not always add a consistent meaning (Allan 2007: 72–4)

1.3 The Arabs

Not long after the foundation of Islam (Hegira AD 622), sophisticated models of language description emerged in the Arabic world largely aimed at exegesis and commentary on the sacred text of the Koran. An early example of Arabic grammar from the eighth century is the work of Sībawayhi, entitled simply *Al Kitab* 'the book'. In Arabic nouns can be marked for number, gender, and definiteness, as well as case. There are three cases: nominative, accusative, and genitive. The Arabic tradition has nothing new to say about cases per se, but it does have some interesting ideas on case assignment, a notion not prominent in the Greek and Roman tradition. The first principle of interest is that nominative and accusative case are assigned to nominals according to the positions they occupy in an underlying order or basic

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 17 of 26 May 16, 2008 18:57

order. In classical Arabic the basic order is clearly VSO. The verb is said to assign nominative to the noun immediately to its right and the accusative to all other nouns not otherwise governed.

(1) *ḍaraba Zaydun ʿAmran.*
hit Zayd ʿAmr
'Zayd hit 'Amr.'

Other word orders are possible, but case assignment is taken to precede reordering.

The other interesting principle is the notion of an abstract case assigner, that is, one devoid of phonetic form. Such an assigner was postulated by some Arabic grammarians to assign the nominative to the subject and predicate of a verbless sentence such as *Zaydun axū-ka* 'Zayd is your brother' (Bohas et al. 1990: 53ff).

1.4 The Modistae

Scholastic philosophy included language in its ambit and scholastic grammarians, or Modistae, as they are often called, made significant contributions to the study of grammar, particularly in the thirteenth and early fourteenth century. They considered the descriptive grammar to be inadequate and sought the higher goal of explanation. They were also interested in the concept of universal grammar. Roger Bacon wrote, 'Grammar is one and the same in substance in all languages, though there may be accidental variation.' *Accidentaliter* 'accidental' needs to be related to *accidentia* 'accidence', i.e. inflection (Robins 1997: 90, 107).

The Modistae introduced a number of ideas into the study of grammar that were new, at least to the western tradition. They distinguished subject (*suppositum*) and predicate (*appositum*) and clear notions of dependence so that, for instance, in *Sōcratēs albus currit bene* 'White Socrates runs well' *Sōcratēs* is suppositum with *albus* its dependent and *currit* appositum with *bene* its dependent. Note in passing the odd, concocted example, a regular feature of Modistae exposition.

The scholastic grammarians were fond of classifying cases on the basis of general properties. One distinction they made was between cases expressing substance-to-action and those expressing substance-to-substance. Substance-to-action was adverbal case, though the vocative was included. Substance-to-substance was adnominal, both attributive and predicative. They were localistic (see above) to some extent. Simon of Dacia (Dacia is Denmark, not Romania) interpreted the Latin genitive as sharing the property of origin with the ablative, and Martin of Dacia took the nominative and genitive to share an origin-like property on the grounds that the action has its origin in the nominative, and in a genitive phrase such as

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 18 of 26 May 16, 2008 18:57

filius Marcī 'the son of Marcus' the possessed has its origin in the possessor. These two cases are opposed to the other four, which are termini, that is, ends in both the literal sense and in the sense of objects or targets. Martin's classification of Latin cases as presented in Marmo (1994: 201f) is as follows. Note that for Martin the ablative is not [+origin] in the relevant sense.

(2)

Origin	adverbal	nominative
	adnominal	genitive
Terminus	adverbal or adnominal	dative 'to', ablative 'from'
	adverbal	accusative, vocative

This is not the only scholastic classification of Latin cases. In fact it is not Martin's only classification, a partly different one is presented in Serbat (1981: 26–7).

1.5 Pāṇini and kāraka theory

Pāṇini's grammar of Sanskrit is known as the Astādhyāyī ('eight books'). Its date is uncertain with estimates ranging from 600 BC to 300 BC. It is a work of great sophistication clearly deriving from a lengthy tradition and it has provided a model for the study of language in greater India down to the present day. As is well known, the 'discovery' of Sanskrit by European scholars in the late eighteenth century and the realization that it is genetically related to Greek and Latin provided great stimulus to comparative-historical linguistics. It has also been claimed that familiarity with Pāṇini and the ancient Indian tradition helped raise the level of sophistication in descriptive linguistics; however, the aspect of Pāṇini's theory that we are concerned with here, namely kāraka theory, has not been influential though discussed by various scholars (Cardona 1976a: 215ff).

Kārakas are semantic relations holding between nouns and verb. The verb is held to be the head of the clause and each nominal dependent is assigned to one of six kārakas. These kārakas are listed in (3) with their common translations. The term 'object' is unfortunate since it suggests a purely syntactic relation. A term like patient would have been more appropriate.

(3)

kartṛ	agent
karman	object
karaṇa	instrument
saṃpradāna	destination
apādāna	source
adhikaraṇa	locus

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 19 of 26 May 16, 2008 18:57

Table 1.2. Cases and kārakas

nominative	*kartṛ*	agent
accusative	*karman*	object
instrumental	*karaṇa*	instrument
dative	*saṃpradāna*	destination
ablative	*apādāna*	source
locative	*adhikaraṇa*	locus

To appreciate the significance of kārakas it is useful to compare them with the case system of Sanskrit, since the cases are obviously the main means of expressing kārakas. In Sanskrit there are eight cases, all of which are distinguished in the singular of masculine *a*-stems. The vocative does not mark a dependent of the verb and therefore a vocative-marked noun bears no kāraka. The genitive is held to be adnominal and so no kāraka is assigned to genitive-marked nouns. However, the genitive does have one adverbal function, namely to mark the complement of a small number of verbs as in: *mātuḥ smara-ti* (mother.GEN remember-3SG) 'S/he remembers his/her mother' (Filliozat 1988: 86). With the other six cases there are regular associations of case and kāraka as shown in Table 1.2.

This list is, however, misleading insofar as it suggests a one-for-one correspondence between case and kāraka and in that it suggests the nominative expresses a kāraka. A consideration of an active sentence and its passive counterpart demonstrates how the one-for-one association needs qualification.

(4) a. *Devadatta odana-m paca-ti*
Devadatta.NOM rice-ACC cook-3SG
'Devadatta cooks the rice.'
b. *Devadatt-ena pac-ya-te odana-ḥ*
Devadatta-INST cook-PASS-3SG rice-NOM
'The rice has been cooked by Devadatta.'

The person marking on a finite verb is held to encode a kāraka, which in the active is the *kartṛ*. The nominative is held to encode only the denotation of the person marker on the verb. Since the nominative-marked noun refers to the same denotatum as the person marker in the verb (*-ti* in (4a)), it names the referent holding the kāraka of *kartṛ* (agent). In the active the accusative encodes the *karman*. In the passive the person marker on the verb is interpreted as encoding the *karman*. The nominative in cross-reference with this person marker (*-te* in (4b)) names the *karman*. The instrumental encodes the *kartṛ*. A consideration of the passive then demonstrates that there is no one-for-one correspondence between case and kāraka and that the nominative does not encode a kāraka. However, Pāṇini's treatment of the passive can give the misleading impression that kārakas are like the deep cases or semantic roles of modern theories, which tend to remain constant under

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 20 of 26 May 16, 2008 18:57

paraphrase. Consider now the following pair and their interpretation (Cardona 1976b: 22–4):

(5) a. *Bhūmāv ās-te Devadatta-ḥ*
ground.LOC sit-3SG Devadatta-NOM
'Devadatta is sitting on the ground.'
b. *Bhūmi-m adhy-ās-te Devadatta-ḥ*
ground-ACC on-sit-3SG Devadatta-NOM
'Devadatta is sitting on the ground.'

In (5a) *bhūmi* is encoded in the locative and interpreted as expressing *adhikaraṇa* (locus). In (5b) the verb *ās* 'to sit' is combined with *adhi* 'on' to yield a transitive verb *adhyās-*. This is analogous to a Latin verb like *īn-stāre* 'to sit on' and is effectively a kind of applicative, where the location can take on extra senses of affectedness. *Bhūmi* becomes the accusative object of *adhyās-* and is interpreted as expressing *karman*. There are a number of examples like this available that show Pāṇini distinguished the linguistic encoding of events from objective reality, that is, he allowed for more than one semantic interpretation of an event.

1.6 The Recent Past

The dominant model for describing case systems in the Renaissance and right down to the present day has been the traditional method of listing a number of meanings and functions for each case. However, in the early nineteenth century the notion of cases having a single, abstract meaning can be found in the work of linguists such as Rask, Bopp, and Wüllner. Not surprisingly this notion of *Gesamtbedeutung* 'generalized meaning' becomes prominent in the structuralist period of the twentieth century. The leading proponents were Hjelmslev and Jakobson, both of whom published major works on case in the mid-thirties.

1.6.1 Hjelmslev

For Hjelmslev, a case signified 'a single abstract notion from which one can deduce the concrete uses' (Hjelmslev 1935: 85). The abstract meaning of a case could only be determined by looking at the oppositions within the whole case system. Hjelmslev was a localist, that is, he characterized the meaning of all cases in local terms, not just the local cases.

To see how this works, let us look at Hjelmslev's interpretation of Greenlandic Eskimo (Hjelmslev 1937: 65–75). In this language there are the following

Table 1.3. Hjelmslev's case system of Greenlandic Eskimo

	[+from, –to]	[+to, –from]	[+from, +to]	[–from, –to]
[–coherent]	ergative	equative	instrumental	nominative
[+coherent]	ablative	allative	perlative	locative

cases: nominative, ergative-genitive, equative (or predicative), instrumental, ablative, allative, locative, and a 'through' case, perlative or prosecutive. The last four of these cases are clearly local, and Hjelmslev captures them in terms of 'to' (*rapprochement*) and 'from' (*eloignement*) as follows:

ablative	from
allative	to
locative	neither 'from' nor 'to'
perlative	both 'from' and 'to'

The second dimension in Hjelmslev's system is degree of intimacy: **coherence** (contact or penetration) versus **incoherence** (proximity) (1935: 128). This is another local notion and Hjelmslev illustrates the distinction with prepositions: *She went **into** the building* (coherent); *she went **to** the building* (incoherent); *she went **between** the buildings* (indifferent to coherence). With Greenlandic Eskimo Hjelmslev uses this dimension to capture the distinction between the local cases and the syntactic cases. He aligns the ergative-genitive with the ablative (the action proceeds from the agent) and the equative with the allative (*A is equal to B*). The nominative which is pivotal and which does not involve either the notion of 'to' or of 'from' is matched with the locative, and the instrumental with the perlative, an analogy which can be rationalized in terms of the action passing through the instrument to the patient. The tabloid summary is presented in Table 1.3 (after Hjelmslev 1937: 74).

1.6.2 Jakobson

Jakobson's approach to determining the meaning of cases echoes Hjelmslev's. He distinguished between the invariant **intensional** meaning of a case and its syntactically and/or lexically conditioned variants, which make up the **extension** of the case. The *Gesamtbedeutung* or aggregate meaning of a case is independent of the environment and cannot be determined from the individual meanings (*Sonderbedeutungen*) nor from the principal meaning (*Hauptbedeutung*). Cases are correlative and take their value from their relation to other cases in a system of oppositions (Jakobson 1936 [1971: 35–6]).

Jakobson applied his theory to the case system of Russian. The cases are: nominative, accusative, dative, instrumental, genitive, and locative, plus two rather

Table 1.4. Russian case system in terms of Jakobson's features

	Marginal	Quantifying	Ascriptive
Nominative	–	–	–
Accusative	–	–	+
Genitive I	–	+	+
Genitive II	–	+	–
Locative II	+	+	–
Locative I	+	+	+
Dative	+	–	+
Instrumental	+	–	–

restricted cases: genitive II or partitive and locative II. Jakobson takes the nominative to be unmarked. Opposed to it is the accusative, which is always subordinated to it and which signals direction or goal. The instrumental and dative are opposed to the nominative and accusative respectively as marginal or peripheral cases (*Randkasus*) opposed to direct or central cases (*Vollkasus*). The instrumental is aligned with the nominative since it expresses a source of an event, most obviously where it expresses the agent of the passive, and the dative is aligned with the accusative in that both express the goal of an event.

The other four cases, genitive I, genitive II, locative I, and locative II focus 'upon the extent to which the entity takes part in the message' (Jakobson 1958/1971: 179), the locative also being considered marginal. The genitive II (partitive) and locative II are taken to be marked with respect to genitive I and locative I respectively. These groupings are shown in (6) where the marked member of each pair is shown to the right, and the four on the right are marked with respect to the four on the left.

(6) central [nom acc] [gen I gen II]
marginal [inst dat] [loc I loc II]

In his 1958 paper Jakobson presented the system in terms of a cube, with what are features serving as the dimensions: [± direction], [± marginal] and [± quantification] (partial involvement). The feature [± directional] is extended from the accusative and dative to cover genitive I and locative I. It is also referred to as [± ascriptive] since 'directional' is not so appropriate when it is extended to cover genitive and locative. Table 1.4 displays the system as a matrix of distinctive features. I think it would be fair to say that Jakobson's characterization of the Russian cases in terms of features is less than perspicuous. It is not clear that 'partial involvement' is appropriate for genitive and locative and there is something arbitrary about the introduction of the feature [± directional] or [± ascriptive]. There is no consideration of alternatives and there is less than adequate demonstration of how the feature analysis can be exploited to capture generalizations.

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 23 of 26 May 16, 2008 18:57

1.7 Beyond Europe

During the twentieth century descriptions of some hundreds of languages became available. The expanded data base allowed a better perspective on the part case can play in language. It also demanded an expansion of the terminology for describing case and raised questions about what counts as case.

In languages like Greek and Latin there is fusion of case and number, and to some extent gender, and the boundary between stem and inflection is not always clear. In these circumstances it makes sense to deal in whole words like *anthrōpos* and *anthrōpou*, what we might call case forms. Case systems occur in agglutinative languages, and in these circumstances it makes sense to consider the inflections as morphemes and to talk of roots or bases, stems and **case markers**. In Turkish, for instance, number marking occurs between the root and the case inflection as in *elma-lar-ı* (apple-PLURAL-ACC).

As more languages are studied, more cases are encountered, which need to be named. One of the most prominent is the ergative, a case for the agent of a transitive verb opposed to another case, almost always unmarked, for the subject of an intransitive predicate and the patient of a transitive verb. This is generally dubbed the absolutive. Another case common in languages of Eastern Europe whether Indo-European or Finno-Ugric is partitive for a partly affected patient.

A question that arises when we look at a wide range of languages is what counts as a case. In Ancient Greek and Latin there was a single small system of case inflection plus a larger system of prepositions. However, many languages do not have these two systems for expressing what we might call case functions.

In languages like Japanese there is a single system of postpositions that expresses grammatical relations such as direct object as with *o* in (7), topic as with *wa*, and local and other relations as with *de*.

(7) *Watasi wa isya o denwa de yobimasu.*
I topic doctor obj phone by call
'I'm calling the doctor by phone.'

The Japanese system in effect neutralizes the distinction between case marker and adposition in favour of a large system of adpositions.

It should be noted that in any language there will be a score or more local relations such as 'over', 'under', 'through', etc. to be expressed. In northeastern Caucasian languages these local notions are expressed by a system of suffixes, which are followed by another system expressing location, destination, and source. In Tabasaran, for instance, there are four non-local cases: absolutive *-ø*, ergative *-i*, genitive *-i-n* and dative *-i-z*. The ergative serves as a stem for all cases other than the absolutive. There is a system of seven or eight orientation markers (the number depends on dialect) such as *-q* 'behind', *-kk* 'under', *-h* 'near, in front of', and *-k* 'on a

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 24 of 26 May 16, 2008 18:57

vertical surface'. These may be followed by *-na* 'allative' or *-an* 'ablative'. These may be followed in turn by *-di*, which indicates a less specific location or direction; with the allative and ablative it means 'in the direction of'. The following examples are adapted from Comrie and Polinsky (1998: 96ff).

(8) *cal-i-q* 'behind the wall'
cal-i-q-na 'to behind the wall'
cal-i-q-an 'from behind the wall'
cal-i-q-na-di 'towards behind the wall'
cal-i-q-an-di 'from the direction of behind the wall'

The system in Tabasaran and other northern Caucasian languages is agglutinative, but some writers have taken sequences of markers to be single case markers, which has lead to claims that these languages have very large case inventories (Comrie and Polinsky 1998). Where orientation markers and case markers proper fuse, we need to take each fused form as a case marker. This has happened in Finnish where forms for 'in' and 'on' have partly fused with the locative (essive), partitive, and translative *-ksi* to yield nine local cases.

In a variety of languages case markers do not all belong to the same system. In some languages a noun bearing an adnominal genitive can take further case marking as in Old Georgian: *sarel-ita man-isa-jta* (name-INST father-GEN-INST) 'with father's name'. In Quechua the same situation arises where a head noun has been elliptically deleted (Weber 1989: 254).

(9) a. *Hwan-pa wasi-n-ta rika-a*
John-GEN house-3SG-ACC see-1SG
'I see John's house.'
b. *Hwan-pa-ta rika-a*
John-GEN-ACC see-ISG
'I see John's (house).'

In most Australian languages there are suffixes for 'having' and 'lacking' that can be considered case suffixes in that they mark the relationship of a dependent noun to another noun as in Kalkutungu *marapai pirlapirla-yan* (woman baby-having) 'the woman with the baby' or 'The woman has a baby'. Such a suffix can be followed by an adnominal genitive as in *pungkuarri pirlapirla-yan-ku* (bag baby-having-GEN) 'bag of the one with the baby', and an adverbal case such as the ablative as in *pungkuarri-thingu pirlapirla-yan-ku-wa-thingu* (bag-ABL baby-having-GEN-LIG-ABL) 'from the bag of the one with the baby' where *-wa-* is a ligative separating successive suffixes.

In Greek and Latin the case system is used for core functions (subject and object) and others such as marking location or instrument. In many languages case marking, whether by suffix or postposition, is used only for non-core functions. With core functions such as subject and object a system of pronominal markers is

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 25 of 26 May 16, 2008 18:57

used, usually on the verb. These may stand for subject or object or cross-reference them as with the object in (10) from the northern Australian language Limilngan (Harvey 2001: 73).

(10) *Limin biyal ngugun=lagarni l-a-na=gi.*
keel.snake water-LOC it-I-see-PAST
'I saw a keel snake in the water.'

As can be seen there is a case marker for locative, but no case marking for the noun object, nor would there be for the subject pronoun, if one had been included. Some writers use case labels for cross-referencing pronominal markers, but since these markers express a particular grammatical relation, it is preferable to call them subject markers and object markers.

With the availability of more data on case systems, cross-language generalizations began to emerge. One of the most robust is Silverstein's demonstration that the distribution of partial ergative marking and partial accusative marking defined a hierarchy of lexical content (Silverstein 1976). The hierarchy is as follows:

(11) 1st person (speaker)
2nd person (addressee)
3rd person pronoun
personal name, kin term
human
animate
inanimate

Partial accusative marking will always run from the top of the hierarchy covering a continuous segment. Partial ergative marking on the other hand will always run from the bottom of the hierarchy. A language may have accusative marking only on first and second person pronouns as in many Australian languages, only on personal pronouns as in English, or only on pronouns, kin terms, and personal names, and so on. All the positions on the hierarchy can be defined on the basis of the distribution of accusative marking, i.e. seven different cut-off points are attested. Ergative case marking tends to be found on all nouns or all nouns and third person pronouns. In languages with a mixture of ergative and accusative marking, the two may overlap in the middle of the hierarchy giving a three-way nominative-ergative-accusative case-marking opposition for some category such as third person pronouns.

In Greek and Latin each verb has a particular case frame. Most two-place verbs are transitive and take a nominative subject and an accusative object, a few take complements in non-core cases. A small number alternate. In Latin *moderor* 'to restrain' takes an accusative complement where the patient is external to the subject: *moderor equum meum* 'I restrain my horse'. It takes a dative complement where the reference is to self-restraint: *moderor ōrātiōnī meae* 'I moderate my speech'. In some

978–0–19–920647–6 02-Spencer-ch01 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 26 of 26 May 16, 2008 18:57

languages there are detransitive verbal derivations yielding intransitive verbs from transitive or there are causative or applicative derivations yielding transitive verbs from intransitive or ditransitive from transitive. In such languages a verbal root can appear in a number of case frames.

Further information on traditional approaches to case and on case data can be found in Blake 2001, and a more recent coverage including modern theories of case can be found in Butt 2006.

CHAPTER 2

MODERN APPROACHES TO CASE

AN OVERVIEW

MIRIAM BUTT

2.1 Introduction

As the range of topics collected in this volume demonstrates, case is a complex phenomenon: in many languages case marking takes on functions that go beyond the purely *structural* role of helping to identify the grammatical relations (subject, object, indirect object, etc.) of a sentence. For example, case marking is used to signal differences in agency, animacy, definiteness/specificity, existence/persistence of an object and is implicated within the domain of tense/aspect in terms of signalling telicity or boundedness. Case furthermore interacts with discourse functions such as topicality or focus (Japanese is a well known example, see Chapter 54) and can express modalities such as obligation vs. desire. Despite the fact that these diverse functions of case have been observed cross-linguistically for quite some time, very few modern approaches to case are currently able to present a coherent analysis of the wide and varied spectrum of functions associated with case.

As the vast range of the contributions in this handbook show, any theory that sought to provide a complete account of case cross-linguistically would not only

need to be in control of a huge amount of data, but also possess a deep understanding of a very varied collection of theoretical approaches. In my opinion, the various perspectives on case offered up by each of the approaches all contain valid insights – only a combination of these valid insights can ultimately result in a complete theory of case.

This chapter cannot hope to do justice to all of the existing modern approaches to case, therefore cannot offer up any such unified theory of case and so instead will concentrate on identifying some important general lines of inquiry that have emerged. Section 2.2 first looks at some central statements about the classification of languages according to case distribution. Sections 2.3 and 2.4 chart some of the leading ideas that have influenced many modern approaches to case. Section 2.5 offers a concluding discussion.

2.2 Language Classification

The twentieth century has been able to establish quite a few useful and interesting generalizations on the distribution of case cross-linguistically.[1] For example, there is the realization that free word order languages tend to make use of case marking (but the converse does not hold: not all heavily case-marked languages necessarily have free word order). Another example of a generalization that has become standard knowledge within general linguistics is Silverstein's (1976) proposal of an implicational preference hierarchy for the appearance of ergative case on subjects, known as the *NP-hierarchy* (see Chapters 1, 37). But perhaps the most influential idea has been the proposed classification of languages into types based on the case marking of their core arguments (see Chapter 45).

2.2.1 Fillmore's Case Relations

A systematic proposal for the cross-linguistic classification of languages on the basis of case was first articulated by Fillmore (1968), who suggested that *case roles* or *case frames* should be acknowledged as one of the common universal bases of language. Under his view, the basic part of a sentence contains a proposition P which contains a tenseless set of relationships involving verbs and nouns. These are *case relationships* (C), as in (1).

(1) $P + V + C_1 + \ldots + C_n$

[1] For a collection of sample generalizations as well as possible counterexamples, see e.g. the *Universals Archive* at http://ling.uni-konstanz.de/pages/proj/sprachbau.htm.

978–0–19–920647–6 03-Spencer-ch02 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 29 of 43 May 16, 2008 18:57

Fillmore argued that the case relationships needed for cross-linguistic analysis include at the very least: *Agentive*, *Instrumental*, *Dative*, *Factitive*, *Locative*, *Objective*. A first cut at a definition is shown in (2) (Fillmore 1968: 24–5).

(2) **Agentive** (**A**) The case of the typically animate perceived instigator of the action identified by the verb.
Instrumental (**I**) The case of the inanimate force or object causally involved in the action or state identified by the verb.
Dative (**D**) The case of the animate being affected by the state or action identified by the verb.
Factitive (**F**) The case of the object or being resulting from the action or state identified by the verb, or understood as part of the meaning of the verb.
Locative (**L**) The case which identifies the location or spatial orientation of the state or action identified by the verb.
Objective (**O**) The semantically most neutral case, the case of anything representable by a noun whose role in the action or state identified by the verb is identified by the semantic interpretation of the verb itself; conceivably the concept should be limited to things which are affected by the action or state identified by the verb. The term is not to be confused with the notion of direct object, nor with the name of the surface case synonymous with accusative.

Given the basic formula in (1), languages are predicted to contain sets of formulas as in (3), which correspond to the basic kinds of sentences. An intransitive clause, for example, might consist of a verb and an agentive or objective case relation, depending on how 'active' the subject was. Note that this distinction is very similar to the one that was later put forward under the *Unaccusative Hypothesis* (see section 2.3.4).

(3)
V + A	(intransitive, active subject)
V + O	(intransitive, inactive subjec)
V + O + A	(transitive)
V + O + D + A	(ditransitive)

Note that the underlying order of case relations as specified by (3) does not match the surface order for English (or, indeed, most languages of the world). Fillmore's idea was that case relations represented a *deep structure* that would then be made to correspond to the surface string (or structure) via a series of transformations of the type available in Transformational Grammar (TG, Chomsky 1957) at the time. Similarly, the case relations abstracted away from the precise realization of case marking – the case roles could be realized by nothing or a preposition (as in English), inflectional case marking, or by some other device.

Fillmore's original ideas were the subject of some controversy. A large part of the discussion revolved around the definitions of the case roles, which turned out to be too vague to be really useful. Fillmore (1977) presents a revision and further

explication of his original ideas, but the notion of *thematic*, *semantic*, or *argument roles* that is now part of general linguistic theory (see section 2.3) owes more to the formulations of Gruber (1965) and Jackendoff (1972), which were developing in parallel.

However, Fillmore's original ideas influenced quite a lot of subsequent syntactic theorizing.[2] As already mentioned, just one of those ideas involved using his abstract case relations to help classify languages by the case marking patterns they display.

2.2.2 Current standard formulation

Fillmore took just two basic case roles to be relevant for the classification of languages: A (agentive) and O (objective). In looking at how languages mark these two case roles in the basic clause types associated with them (V+A, V+O, V+O+A), Fillmore came up with a classification of language types that led him to recognize five different types of case systems. However, of the entire classification, only a two-way opposition between what are now called *ergative* and *accusative* languages became standardly recognized.[3] In particular, the classification shown in (4) came to be the standard way of illustrating the distinction (Dixon 1994).

(4)

nominative	A	ergative
	S	absolutive/nominative
accusative	O	

A = transitive subject (Agent), S = intransitive subject,
O = transitive object

As shown in (4), *ergative* languages are classified as those which tend to mark the agent with an ergative and the objective/object (O) with a nominative or absolutive. This nominative or absolutive tends to be exactly the same marker that is used for subjects of intransitives. In contrast, languages are classified as *accusative* when they distinguish the object via an accusative case and generally group subjects together by marking subjects of both transitive and intransitive clauses consistently with the nominative (English and German would be examples).

[2] Fillmore saw himself as proposing a variant of TG; however, his *Case Grammar* turned out to be less of a variant of TG and more of an independent theoretical persuasion which shares some fundamental ideas with Tesnière's (1959) *Dependency Grammar*. Today's *Localist Case Grammar* (Chapter 8) owes much to Fillmore.

[3] In addition to this distinction, Fillmore also allowed for languages like Yana, which treat pronominal As and Os in all sentence types alike; Dakota, which distinguishes all As from all Os; and Takelma, which has separate forms for almost everything.

2.2.3 Evaluation

The attentive reader will have noted that the standard method of classifying languages involves a 'new' notion: S. The addition of the notion S to Fillmore's original case roles allows a clear distinction between intransitive subjects and objects. This in turn allows for a cleaner statement of the generalizations differentiating between accusative-type and ergative-type language: accusative languages distinguish O, ergative languages distinguish A. However, it also obscures the difference between different types of intransitive clauses that was built into Fillmore's system and that indeed mirrors a cross-linguistic division of intransitive verbs into at least unaccusative vs. unergative verbs (see section 2.3.4).

This distinction has been built back into the current system by acknowledging an 'active' type of language which distinguishes unergative verbs from unaccusative ones via case marking (some examples are Georgian and Urdu). Other distinctions have also been introduced into the system, as it is clear that a complete cross-linguistic typology needs to take into account more than the now standard two-way distinction between ergative and accusative languages and, indeed, recognize more than Fillmore's original five distinct systems. For more information on typologies of case systems and the current state of the art, see Chapter 45.

One issue that so far has not been the subject of much discussion, however, is Fillmore's original assumption that the pertinent patterns for the classification of languages are: 1) the case of the agentive subject vs. the intransitive subject; 2) the canonical transitive frames (V+O+A) and the intransitive ones. This very fundamental assumption has far reaching effects, and should be discussed explicitly.

Fillmore's original assumption excludes a consideration of indirect objects. But what justifies this assumption? Some languages use case to identify indirect objects (generally with the dative case), some do not – why should this not provide a useful basis for classification into language types? Furthermore, why should one not be forced to take into account case marking patterns in clauses that are not canonically transitive (cf. Hopper and Thompson 1980, also Chapter 23)?

The focus on canonically transitive clauses (V+O+A in Fillmore's terms) has meant an exclusive focus on nominative/absolutive, accusative, and ergative case marking. Other case markers such as dative, instrumental, or genitive are automatically excluded. However, many languages use datives or genitives to mark subjects of so-called *experiencer* or *psych* verbs as in (5); other languages do not. Why should this not also be a relevant factor?

(5) a. *Mér batnaði kvefið.*
I.DAT recovered the.cold.NOM
'I recovered from the cold.' (Svenonius 2002: 205) Icelandic

b. *nadya=ko dar aya*
Nadya=DAT fear come.PRF.M.SG
'Nadya got scared.' (lit. 'fear came to Nadya') Urdu

978–0–19–920647–6 03-Spencer-ch02 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 32 of 43 May 16, 2008 18:57

Furthermore, many languages show case alternations, either on subjects or on objects.[4] The examples from Georgian in (6) and from Urdu in (7) illustrate an ergative–dative subject alternation.

(6) a. ***nino-m*** *Ceril-i* *daCera*.
Nino-ERG letter-NOM wrote-3S;3O
'Nino wrote a letter.' Georgian

b. (*turme*) ***nino-s*** *Ceril-i* *dauCeria*.
apparently Nino-DAT letter-NOM wrote-3SGS;3O
'Apparently Nino wrote a letter.' Georgian

(7) a. ***nadya=ne*** *zu* *ja-na* *hɛ*
Nadya.F.SG=ERG ZOO.M.SG.OBL go-INF.M.SG be.PRES.3.SG
'Nadya wants to go to the zoo.' Urdu

b. ***nadya=ko*** *zu* *ja-na* *hɛ*
Nadya.F.SG=DAT ZOO.M.SG.OBL go-INF.M.SG be.PRES.3.SG
'Nadya has to/wants to go to the zoo.' Urdu

The alternations are regular and have generalizable semantic import. In Georgian, the use of the dative instead of the ergative case goes hand in hand with a meaning of 'apparently'. In Urdu, the dative is the unmarked case and can express a modality of either desire or obligation, depending on the context, whereas the ergative unambiguously signals desire.

The existence of such alternations is unexpected given a purely structural typology of languages in which only A, S, O, and nominative/absolutive vs. ergative vs. accusative case are taken to play a role. Furthermore, not all ergative languages display such a dative–ergative alternation. It would therefore seem to be typologically useful to establish which languages allow for such an alternation (and why, of course). However, despite the fact that reports on the existence of the above alternations are not new (e.g. Harris 1981, Butt and King 1991), to my knowledge they have not triggered a typological investigation that takes such alternations into account.

Finally, note that although some of the fundamental assumptions behind the current practice of language classification can and should be questioned, the idea of a systematic classification of languages via case frames was very innovative in that it differs markedly from previous approaches, which tended to concentrate on trying to find a unifying semantics for individual case morphemes, as in the work of Jakobson (1936) or the localist tradition (e.g. Hjelmslev 1935), cf. Chapter 1.

[4] *Differential Case Marking*, as it has come to be known, has been a recent focus of optimality-theoretic approaches, see Chapter 6.

978–0–19–920647–6 03-Spencer-ch02 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 33 of 43 May 16, 2008 18:57

2.3 Leading ideas: (morpho)syntax and semantics

A comparative look at the underlying ideas contained in syntactic theories from antiquity to modern times (see Butt 2006) identifies several major lines of enquiry that have been postulated to help understand the distribution of case. Some of these include the notion of semantic roles (section 2.3.1), lexical decomposition (section 2.3.2), proto-roles (section 2.3.3), grammatical relations (section 2.3.4) and linking (section 2.3.5).

Since the idea of semantic roles has already been introduced with respect to Fillmore in section 2.2, we begin with this topic.

2.3.1 Semantic roles

Pāṇini's *Kāraka Theory* (see Chapter 1) was the first to propose a systematic relationship between semantic roles such as agent, object of desire/goal, instrument, etc. and overt case marking. A quick look at South Asian languages, for example, shows that agents are generally marked with an ergative or instrumental, goals/experiencers with a dative, patients with an accusative, etc. Indeed, cross-linguistic data confirms that there is a clear semantic basis at the heart of a great deal of the distribution of cases (see Chapter 11 and Butt 2006 for examples and more discussion).

Most theories of case today assume that predicates (verbs, nouns, prepositions, and also adjectives) come with some kind of underlying specification as to their *argument structure*, that is, a specification as to the number and semantic type of participant roles involved (see Chapter 17 for details),[5] and that this information is relevant for capturing case marking patterns across languages. Where theories tend to differ is on questions of representation and the precise interaction with other modules of the grammar (see Levin and Rappaport Hovav 2005 for a comprehensive discussion).

Generally, at least the following thematic roles are acknowledged: *agent*, *goal/experiencer*, *theme/patient*, *instrument*, *location*. However, across theories there is a huge amount of dissatisfaction with these role labels – this dissatisfaction arose when Fillmore (1968) first proposed his ideas and came to a head in the summary of the state-of-the-art provided by Dowty (1991). The problem is that while the labels are utterly intuitive and therefore quite useful at some level of description, it is very difficult to put them to practical use because the definitions provided to date are simply too vague.

[5] A current exception is work by Marantz (1997) and Borer (2005), who attempt to eschew an argument structure altogether.

2.3.2 Semantic primitives

The solutions that have been pursued come in various flavours. One popular solution, going back to the efforts of Generative Semantics in the 1970s, has been to identify *semantic primitives* which are relevant for understanding the lexical semantics of predication and the number and type of participants involved. A prominent researcher representing this direction of enquiry is Jackendoff (e.g. Jackendoff 1972, 1976, 1987, 1990).

Jackendoff works mainly on English, so he does not tend to figure centrally in discussions of theories of case; however, his framework is in principle eminently suitable for describing case relations. Consider, for example, the lexical entry for 'give' in (8) (for a slightly more detailed discussion see Butt 2006, for an application of this system to case marking in Urdu, see Butt 1995).

(8) CAUSE([α], GO$_{Poss}$ ([β], TO[ ]))
AFF([ ]$^{\alpha}$, [ ]$^{\beta}$)

In this lexical entry (the Lexical Conceptual Structure or LCS in Jackendoff's terms) the event denoted by the English verb *give* is lexically decomposed into several primitives, namely CAUSE, GO, and TO. The representation expresses that there is an event in which a participant (first argument of CAUSE) caused something (first argument of GO) to go to another participant (argument of TO). The verb *give* thus has exactly three arguments which need to be linked into the syntax.

Most readers will be aware of the English *dative alternation* whereby the arguments of 'give' can be realized via a double object construction: *The monkey gave the dog a bone.* This possibility is dealt with quite elegantly in Jackendoff (1990) via the introduction of an *Action Tier*, which codes an AFF(ectedness) relation.

In (8), the two participants on the Action Tier are the causer of the event (coindexed with an α) and the theme/patient (coindexed with a β). The Action Tier privileges two event participants, the actor and the patient/beneficiary of the action. These two arguments are therefore realized as the subject and the object of the clause, respectively. Note that the postulation of an Action Tier nicely models the much cited insight that the typical verbal event is transitive and involves two participants, an actor and a thing or person acted upon (cf. Hopper and Thompson's 1980 notion of transitivity, Chapter 23), but does this without a parallel restriction in focus on just canonical arguments.

Given the AFF relation, the double object realization of the arguments of 'give' is accounted for by a version of (8) in which the argument of TO and the second argument of AFF are coindexed (9). This signals that the first argument of GO, the beneficiary, should be interpreted as the affected participant in (9) and therefore become the direct object in syntax.

(9) CAUSE([α], GO$_{Poss}$ ([ ], TO[β]))
AFF([ ]$^{\alpha}$, [ ]$^{\beta}$)

Within Jackendoff's system, there is thus no direct reference to thematic roles. Rather, they could be seen as emerging from the structural positions in the lexical semantic representation. And these structural positions in turn can easily be associated with case marking patterns cross-linguistically (i.e. arguments of CAUSE are generally nominative or ergative, arguments of TO are dative or marked with a preposition, etc.).

2.3.3 Proto-roles

A different solution to the problem of unsatisfyingly vague definitions of thematic roles involves the postulation of *macroroles* (Van Valin 1977) or *proto-roles* (Dowty 1991) (cf. Chapter 17). Rather than worrying about whether a participant can indeed be identified as an *agent* when that participant is not volitional, for example, the postulation of proto-roles revolves around the idea that a typical agent or *actor* role is actually a collection of various semantic properties, not all of which need to hold all the time. Two main proto-roles are usually assumed: *Proto-Agent* and *Proto-Patient*. Again, this reflects the insight that the typical verbal event is transitive and involves two participants: an actor and a thing or person acted upon.

Many theories have adopted some version of the proto-role approach. Van Valin was the first to pioneer such an approach in the form of Role and Reference Grammar (RRG; Chapter 7) and macro roles play a central role in determining case marking. At the same time, RRG is also a theory that makes a systematic use of semantic primitives at the level of lexical semantics. The information coming from these semantic primitives is bundled and flows into the determination of macro role properties.

In principle, the RRG approach is the most consistent as it first works with fine-grained lexical semantic properties, bundles those properties together in terms of macroroles, and then includes a consideration of syntactic factors in the final determination of the case marking of arguments. However, not all theories have adopted this approach.

Some versions of Linking Theory (section 2.3.5) within Lexical-Functional Grammar (LFG; Chapter 4) have adopted proto-roles, but there is no standard/consistent method of representing the lexical semantic factors that contribute towards the determination of proto-roles (e.g. causation, animacy, volitionality, affectedness), though there is an acknowledgement that these things must of course be represented somewhere.

In contrast, Goverment-Binding (GB)/Minimalism (Chapter 3) has not incorporated proto-roles at all, but instead has concentrated on finding direct correlations between lexical semantic decomposition and syntactic reflexes. That is, syntactic structure is assumed to mirror underlying lexical semantic relationships very directly. Earlier versions of the theory made a direct reference to thematic role labels

like *agent*, *patient*, *experiencer*, etc. In recent years, there has been an effort to have these labels emerge out of fine-grained syntactic structures that mirror/represent lexical semantic decomposition. A central paper here is Hale and Keyser (1993); a recent effort, which also includes discussions of the distribution of case marking, is Ramchand (2007).

Optimality-Theoretic approaches (Chapter 6) have no standard approach: constraints are formulated with respect to thematic role labels, semantic primitives, as well as proto-roles.

Finally, Cognitive Grammar (Chapter 9) and Localist Case Grammar (Chapter 8) use thematic role labels directly as an important part of the theory, while Case in Tiers (Chapter 5) eschews the use of thematic roles completely, instead concentrating on a hierarchy of case in conjunction with a hierarchy of grammatical relations.

2.3.4 Grammatical relations, unaccusativity

Notions like the *subject*, *object*, or *indirect object* of a clause are used pretheoretically by most linguists and most linguists carry around with them expectations that subjects should tend to be nominative, objects accusative, and indirect objects dative (note that cross-linguistic data does not actually fulfil these expectations). Only a few theories, however, grant the notion of *grammatical relations* an official status. One of the first theories to do so was Relational Grammar (RG). This section first looks at the role of grammatical relations in RG, then introduces the *Unaccusativity Hypothesis*, which is due to RG, and then surveys the use of grammatical relations in modern approaches.

In searching for a generalization that would be universally relevant for the formulation of a *passivization* rule, Perlmutter and Postal (1983b) concluded that the only satisfactory generalization was one that was stated with respect to grammatical relations: in passives the object of the active clause becomes the subject of the passive and the subject of the active is demoted to an oblique or adjunct.

RG itself did not use labels like *subject* or *object*, but instead introduced a notion of *terms*, which were labelled 1, 2, and 3. Though not stated explicitly, there was some expectation that 1s would be nominative, 2s accusative, and 3s dative. There was also some expectation that 1s generally correspond to subjects, 2s to objects, and 3s to indirect objects. Furthermore, 1s were often placed in correspondence with agents, 2s with patients/themes, and 3s with goals. However, as the broad range of cross-linguistic research done within RG showed, none of these correlations hold exactly.

For example, the Georgian ergative–dative alternation presented above in (6) was analysed as an instance of 3-to-1 advancement. That is, the dative argument begins life underlyingly as a 3 (explaining the dative marking), but ends up on the surface as a 1 (modelling its status as a subject).

978–0–19–920647–6 03-Spencer-ch02 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 37 of 43 May 16, 2008 18:57

RG's postulation of terms thus allowed for a very integrated account of the interaction between lexical semantics, case marking, and grammatical relations. This type of account is not mirrored as such in any of the other modern approaches to case, though many have adopted insights formulated within RG, a prominent example being the Unaccusativity Hypothesis.

Perlmutter (1978) observed that at least two classes of intransitive verbs exist: *unaccusatives* (e.g. *fall*, *melt*, generally non-agentive, verbs of motion or change-of-state) and *unergatives* (e.g. *dance*, *sneeze*, generally agentive, verbs of bodily process, etc.).

In languages like Urdu, this distinction goes along with a difference in case marking. Unergative verbs have an ergative, unaccusatives are unmarked (nominative).[6] In some languages, the subject of unaccusatives even surfaces as an accusative.

(10) a. *nadya* *gɪr-i*
Nadya.F.NOM fall-PRF.F.SG
'Nadya fell.'
b. *nadya=ne* *kʰãs-a*
Nadya.F=ERG cough-PRF.M.SG
'Nadya coughed.'

The analysis proposed by RG assumes an underlying 2 for unaccusatives and an underlying 1 for unergatives. This roughly models underlying agentivity for unergatives and underlying patienthood for unaccusatives. As all sentences are assumed to require subjects (a final 1 in RG), the underlying 2 of unaccusatives also ends up as a 1.

Essentially this analysis has been adopted into theories like LFG, GB/Minimalism, and RRG, whereby each theory differs in its assumption as to how semantic information interacts with the syntax. LFG took seriously the idea that an abstract level of representation – grammatical relations/RG's terms – that is not equivalent to thematic roles or case should be assumed. So LFG posits terms as well, but labels these with the more traditional labels SUBJ, OBJ, etc. Within LFG, unaccusatives are thus coded as having a patient argument at argument-structure that is then mapped to a SUBJ as part of a *Linking Theory*, which places thematic roles, case marking, and grammatical relations in correspondence with one another (see section 2.3.5).

RRG assumes a clear semantic basis for the unaccusative/unergative split (Van Valin 1990) and models the difference in terms of sophisticated lexical semantic representations which result in a difference in macro roles (Actor vs. Undergoer). Whichever the macro role, the single argument of both unaccusative and unergative verbs ends up being the *syntactic pivot* of the clause. That is, RRG has not adopted

[6] The statement that ergative case appears on unergative verbs tends to be confusing, but that is because of an unfortunate history of nomenclature with respect to *unaccusative* and *unergative* verbs (see Pullum 1988) – these terms were coined without taking ergative patterns as in (10) into account.

a notion of grammatical relations, but rather focuses on the idea that in each clause there is one syntactic pivot with respect to which important generalizations, including ones about case marking, can be stated.

The GB/Minimalism analysis is in many ways quite parallel to the original RG one: an unaccusative is analysed as having a V with an underlying complement (patient) that then moves to Spec, IP (or some similar position) where it receives nominative case (i.e. an underlying 2 moves to become a surface 1). However, very much unlike RG, GB/Minimalism in its current form does not include a notion of grammatical relations – the labels *subject*, *object*, etc. are used, but in a pre-theoretical fashion. The theory also does not include a notion of syntactic pivot. Rather, the particular position an argument is situated at within a syntactic tree is considered to be significant and is interpreted in various ways.

Other theories in which grammatical relations do not play a significant role are Cognitive Grammar and Localist Case Grammar. In contrast, the Case in Tiers approach relies heavily on grammatical relations in that it arranges them in a hierarchy and then formulates correspondences between this hierarchy and a *case hierarchy* (Chapter 5).

To sum up: the unaccusative/unergative distinction proposed within RG is now considered to be part of standard linguistic knowledge. However, the idea that grammatical relations or terms play a significant role in helping encode generalizations about case marking (and other morphosyntactic phenomena) has not been adopted across the board.

The next section turns to *Linking*, that is, a formulation of the relationship between thematic roles, case, and grammatical relations (if your theory has them).

2.3.5 Linking

Although RG's generalization for passivization was stated at the level of *terms*, intense work on lexical semantics has since then shifted the analysis to the level of lexical semantics. A sample LFG analysis is sketched in (11).

(11) *pinch* < agent patient >
| |
active SUBJ OBJ
passive SUBJ (suppression of agent argument)

That is, in the active version of *pinch*, both participant arguments of the verb are linked into the syntax. Specifically in LFG, they are linked to grammatical relations. Passive morphology (or auxiliaries) are assumed to have the effect of suppressing the agent so that it is not available for linking to the subject anymore. Instead, the patient is linked to the subject.

978–0–19–920647–6 03-Spencer-ch02 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 39 of 43 May 16, 2008 18:57

The essential problem to be solved, under this world view, then, is how to relate the semantic participants of a predicate to their syntactic expressions. Most *Linking Theories* or *Mapping Theories* assume a many-to-many mapping and attempt to find constraints on and generalizations over these mappings.

Some generalizations are simple: cross-linguistically agents are mostly realized as *subjects*, *patients* as *objects*, and *goals/recipients* as *indirect objects*. Agents also generally are associated with nominative or ergative case, patients with accusative case, and goals with the dative. But this is just part of a larger pattern.

One of the most sophisticated theories of linking has been articulated by Kiparsky (1987, 1988, 1997, 2001), who provides an elegant and complex picture of the interaction between case, agreement, position, argument structure, and grammatical relations. Kiparsky integrates a sophisticated semantic perspective on case as well as drawing on Pāṇini's *Kāraka Theory* (see Chapter 1). Furthermore, Kiparsky identifies case, agreement, and structural position as three *linkers* which all play a role in identifying the grammatical relations (Abstract Case in his terms) of a clause. Languages may make more or less use of each of these encoding strategies, but at least one of the three is bound to be utilized by a language.

Most theories seem to contain some version of Kiparsky's idea of linkers, though none formulate it as explicitly as Kiparsky. The majority syntax view, which is Government-Binding (GB)/Minimalism, acknowledges both structural position and agreement as playing a major role. Case markers themselves, however, are derivative as they are seen as mere *spell-outs*, that is, as pieces of morphophonology that emerge as part of the pronunciation of a particular structure, but do not play a role in determining the structure. A further difference is that GB/Minimalism adheres to the idea of a strict one-to-one mapping between semantic roles and structural position.[7] This idea is generally known as the UTAH (Uniformity of Theta Assignment Hypothesis, Baker 1988: 46). A patient, for example, is assumed to always be accusative, structurally positioned as the complement of the verb (which roughly corresponds to being the object) and to not agree with the verb. Despite numerous difficulties in the face of empirical cross-linguistic patterns, the majority of GB/Minimalism practitioners continue to follow the very strong encoding of the one-to-one mapping assumption, as this is considered to be more theoretically elegant and predictive.

GB/Minimalism is thus the only modern approach which avoids a linking approach. The other theories all assume a many-to-many mapping, though differ on the details and the type of involvement of information across modules of grammar.

LFG, for example, factors in positional information as well as information from case markers. That is, position and case are seen as linkers in Kiparsky's terms (cf. Nordlinger 1998b and this volume, Chapter 4 for an analysis whereby case

[7] As mentioned, grammatical relations per se do not play a role in the theory, but are assumed to be directly related to structural position.

markers constructively contribute to the syntactic analysis of a clause). Agreement, on the other hand, is seen as more of a requirement on structural well-formedness than a truly useful indicator of grammatical relation status. This is because the cross-linguistic pattern of agreement is quite varied and no useful generalizations really seem to be emerging.[8]

Like Kiparsky's approach, RRG views structural position, agreement, and case all as being implicated in the process of linking to syntax. In particular, RRG contains the notion of a *syntactic pivot* which is identified by agreement and often also nominative case. RRG explicitly implicates case as an inherent part of the process of linking, but like GB/Minimalism and unlike LFG, 'assigns' case markers as part of the syntactic analysis, rather than having them carry information of their own.

The other theories surveyed in this book are not linking theories. Approaches within OT could in principle incorporate an explicit theory of linking, but this has not been done in practice. A collection of constraints on relations between thematic roles, semantic properties of the clause or parts of the clause, grammatical relations, and case is formulated, but this collection remains loose and varies from paper to paper.

2.4 Cognition, quirky case, and formal semantics

Most theories agree that case marking, among other things, expresses a relationship between a predicate and its dependents in that the predicate *governs* the form, type, and sometimes also the position of the dependent. Theories differ on the exact type of government involved and how it should be realized, but most theories show similarities in that they consider some types of predication to be more basic or 'core'. Recall that only the patterns V+O+A, V+O, and V+A played a role for Fillmore in terms of language classification. Recall also that proto-roles are couched so that they identify core 'transitive' predications. Other types, like occurrences of

[8] One core assumption that continues to be held quite strongly in some theories (particularly strongly in GB/Minimalism) is that the verb will always agree with the nominative. However, this core assumption is falsified with respect to a number of languages, among them Nepali.

i. *mai=le* *mero* *lugga* *dho-en*
I=ERG PossPron.1M.SG.GEN clothes.M.PI wash-PRF.1.SG
'I washed my clothes.'

Nepali has an ergative marker for agentive subjects. In (i) the first singular subject is ergative, the object is unmarked (usually glossed as nominative in the literature). Verb agreement here is very clearly with the non-nominative subject.

978–0–19–920647–6 03-Spencer-ch02 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 41 of 43 May 16, 2008 18:57

the dative on indirect objects or ergative-dative alternations as illustrated in (6) for Urdu and Georgian tend to be considered deviations from the norm.[9] As already mentioned, in addition to the assumption that only some patterns are 'core', there was also a strong underlying assumption that subjects are nominative (or ergative), objects accusative, and indirect objects dative. This assumption is pre-theoretic (particularly for those theories that do not rely on grammatical relations as part of the theoretical apparatus), but quite strong. The assumption is also quite wrong, as was shown by Zaenen, Maling, and Thráinsson (1985).

In a landmark paper, Zaenen, Maling, and Thráinsson very carefully and thoroughly established that, over and above allowing various types of non-canonical case marking on objects, Icelandic also allowed dative, accusative, and genitive subjects. An example for a dative subject is (12).

(12) *Mér batnaði kvefið.*
I.DAT recovered the.cold.NOM
'I recovered from the cold.' (Svenonius 2002: 205) Icelandic

Although Zaenen et al.'s paper shows that a generalizable regularity holds between types of thematic roles and case markers (goals are generally datives, etc.), they themselves first offered an account for the non-canonical case marking in terms of lexical stipulation. That is, the lexical entries of verbs were analysed as providing information about the types of thematic roles involved and then furthermore stipulating the type of non-canonical case markers involved.

Within GB/Minimalism, their analysis was interpreted to motivate a difference between *structural* (nominative and accusative) and *inherent* or *quirky* (all other kinds of) case. The former is taken to be regular and assigned as part of the syntactic structure of a language. The latter is considered to be associated with the verb's lexical entry and is therefore assumed to be irregular, inexplicable, and not worthy of interest.

However, cross-linguistic data shows that the patterns identified for Icelandic are not wholly unique, but that so-called non-canonical case marking occurs again and again with respect to generalizable semantic factors.[10] In particular, much seems to be due to spatial metaphors, as can be seen quite readily with respect to so-called *experiencer* or *psych verbs*.

The observation is that the subjects of verbs like *fear*, *feel*, *like/love*, *be hungry*, *be cold*, or *recover* as in (12) cross-linguistically tend to deviate from the default subject

[9] Theories differ on how far they see dative indirect objects as deviations from the norm. In GB/Minimalism extra machinery in terms of an extra governor (usually a null verbal head) must be introduced. RRG assigns datives to those roles which are neither Actor nor Undergoer macro roles. RG, in contrast, considers dative indirect objects as part of the basic inventory (term 3), as do the linking theories assumed within LFG, Kiparsky's theory, as well as Wunderlich's Lexical Decomposition Grammar (see Chapter 17).

[10] With respect to Icelandic in particular, some recent work in this direction is e.g. Svenonius 2002, Eythórsson 2002.

marking. Very often this 'non-canonical' case is a dative. Less often, it might be an accusative or a genitive. More precisely, examples such as (13) and (14) suggest that there is some connection to spatial semantics. In both examples, the main verb is a spatial verb. In Bengali, the liking of tea is metaphorically attached to the subject; in Urdu, fear can be seen to have metaphorically come to one. Experiencer subjects can thus be analysed as abstractions over originally spatial configurations (cf. Verma and Mohanan 1990).

(13) *amar* *ʧa* *bʰalo* *lage*
I.GEN tea.NOM good be.attached.PRES
'I like tea.' (Klaiman 1980: 276) Bengali

(14) *mʊ̃jʰe* *dar* *aya*
I.DAT fear come.PRF.M.SG
'I got scared.' (lit. 'fear came to me') Urdu

Further instances of underlying spatial metaphors can be similarly readily identified with respect to other case patterns. *Localistic Case Grammar* (Chapter 8) therefore builds heavily spatial/localistic concepts. Indeed, spatial metaphors and other semantic extensions are so widespread that *Cognitive Grammar* (Chapter 9) sees this as the primary factor underlying the distribution of case marking.

While this theoretical stance is clearly right in a very fundamental way, examples as in (15) and (16) present a challenge. Here, a difference in case marking goes hand-in-hand with a difference in modality. How can spatial metaphors (or other semantic extensions) be used to explain modal semantics in a straightforward manner?

(15) a. ***ami*** *tomake* *cai*
I.NOM you.ACC wants
'I want you.' (Klaiman 1980: 279) Bengali

b. ***amar*** *tomake* *cai*
I.GEN you.ACC wants
'I need you.' (Klaiman 1980: 279) Bengali

(16) a. ***amma*** *kuṭṭiye* *aḍik'k'-aṇam*
mother.NOM child.ACC beat-want
'Mother must beat the child.' Malayalam
(Butt, King, and Varghese 2004)

b. ***ammak'k'ə*** *kuṭṭiye* *aḍik'k'-aṇam*
mother.DAT child.ACC beat-want
'Mother wants to beat the child.' Malayalam
(Butt, King, and Varghese 2004)

There is thus clearly much more work to be done with respect to explaining the connection between case marking and semantic expression.

978–0–19–920647–6 03-Spencer-ch02 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 43 of 43 May 16, 2008 18:57

2.5 A COMPLEX PICTURE

This short overview chapter cannot begin to do justice to all of the modern approaches to case. In particular, there are several other theories that deal with case, but which are not discussed in the handbook. Among these are theories of syntax such as Head-Driven Phrase Structure Grammar (HPSG), Tree Adjoining Grammar (TAG), and Combinatory Categorial Grammar (CCG). With respect to the core ideas sketched in this chapter, these theories can be seen as following the intuitions articulated within GB/Minimalism quite closely, though they are realized quite differently in terms of technology and representations.

In addition to sketching some core ideas and assumptions of modern approaches in this chapter, I have also tried to show that any analysis of case must be prepared to encounter a complex picture of interactions and, in particular, must be prepared for a semantic base underlying many of the patterns.

However, not all of the patterns are semantic in nature. One aspect that has not been dealt with much is the idea of the *Case Tier* (Chapter 5). This approach follows the general assumptions of GB/Minimalism, but additionally proposes an alignment between case markers and a *hierarchy* of grammatical relations. This proposal is based on evidence that languages do organize at least part of their case marking system in terms of purely structural alignment constraints between case and grammatical relations. This idea is contained in very few other approaches, but seems to have pin-pointed yet another piece of the complex case picture, which involves an intricate interaction between morphosyntactic, semantic, and discoursal factors.

CHAPTER 3

CASE IN GB/MINIMALISM*

JONATHAN DAVID BOBALJIK
SUSI WURMBRAND

Case Theory in GB/Minimalism is about the distribution of NPs, not about morphological form per se. In this theory, 'abstract case' plays a central role in being one of the driving forces of movement, uniting a variety of transformations (passive, raising, unaccusative, etc.), and in regulating alternations between overt and unpronounced subjects in non-finite clauses. In the original presentation of Case Theory in Chomsky (1980), abstract Case is related to the morphological property case via the hypothesis that the formal features that regulate the syntactic distribution of NPs are the same features that are overtly realized as case morphology in some languages.[1] As the theory developed, and in particular after prominent attention was given to quirky case in Icelandic and ergative case systems, the connection between Case (a formal feature underlying syntactic licensing of NPs) and case (the morphological category) became more tenuous, though the connection between the two is still a live topic of inquiry, with views spanning the spectrum of possibilities.

* We thank Željko Bošković and an anonymous reviewer for their comments on a draft of this chapter. Space limitations have kept us from addressing all of their comments, and from doing justice to various topics within Case Theory.

[1] We write 'Case' for the abstract theoretical entity in GB and Minimalism, and 'case' for the traditional morphosyntactic notion. As discussed below, the two notions were originally held to be directly related, but the relationship has grown more abstract.

3.1 Case Theory in GB

3.1.1 The Case Filter

Case Theory is first proposed in the defining works of the GB framework (Chomsky 1980, 1981) as a solution to the puzzling distribution of lexical (i.e. phonologically overt) NP subjects of infinitival clauses in English, as illustrated in (1). In general, the subject of an infinitive must not be an overt NP (1a,e), but this restriction is lifted when the infinitival clause is the complement of a particular class of matrix verbs, such as *believe* (1b), or when the infinitival clause contains the prepositional complementizer *for* (1c,d). Where an overt lexical NP subject is prohibited, the subject of the infinitive is assumed to be the silent pronominal element PRO, the interpretation of which is determined by Control Theory.

(1) a. Leo decided [(*Lina/himself) to leave].
b. Leo believed [Lina to be a genius].
c. Leo decided [for Lina to leave].
d. For Leo to win would be great.
e. *Leo to win would be great.

Prior to the advent of Case Theory, this distribution fell under the purview of the **NP-to-VP* filter of Chomsky and Lasnik (1977), given in (2).

(2) * [α NP to VP], unless α is adjacent to and in the domain of Verb or *for* ([-N])

The major result of Case Theory is the deduction of (2), itself little more than a summary of the facts in (1), from assumptions that are argued to be independently necessary, along with one new assumption, of far broader generality than the construction-specific filter in (2). This new assumption, the key element of Case Theory, is the proposal that all lexical NPs (i.e. NPs other than PRO or NP-trace) require Case, even in Modern English, where the morphological exponence of case is limited to the pronominal system. This proposal (which Chomsky attributes to Jean-Roger Vergnaud) is formalized as the Case Filter, given in one version in (3).[2]

(3) *NP if NP has phonetic content and has no Case (Chomsky 1981: 49)

Given the Case Filter, the distribution in (1) may be largely deduced from independently motivated rules of Case assignment. A rather rudimentary statement of Case assignment for English (and similar nominative-accusative languages) is given in (4):

[2] Note that the Case Filter does not ban a phonetically empty element with Case. Thus in early GB implementations, *wh*-trace is taken to be Case-marked, and a moved *wh*-phrase satisfies the Case Filter via its trace. In later implementations, the Case Filter applies to A-chains.

(4) a. subject of tensed clause: nominative
b. object of verb: accusative
c. object of preposition: accusative (or oblique)

The (descriptive) content of (4) is a necessary part of any grammatical description of English and summarizes the observed basic distributional facts for elements that bear overt case inflection. Absent from (4) is any reference to the subject of a non-finite clause. Given the Case Filter, the absence of case assignment rules applying to the subjects of infinitives translates into the exclusion of lexical NPs from this position. In addition, the 'unless' clause of (2) begins to make sense when viewed from the perspective of Case. That is to say, verbs and prepositions have the distinctive characteristic of being (accusative) case assigners, and thus the disjunctive environment stipulated in the 'unless' clause is none other than the domain of accusative case assignment.[3] What (2) amounts to is that the subject of an infinitive may not be lexical, unless it is in the domain of a case assigner. This proposition is straightforward for the NP following *for* in (1c,d) but requires some additional assumptions for the NP in the infinitival complement of *believe* (1b), to which we now turn.

The contrast in (1a–b) shows that there is some difference between the class of verbs represented by *decide* and that represented by *believe*. The Case Filter provides an account of this contrast, if what is special about the *believe* class is that they permit Case assignment across a non-finite clause boundary.[4] That the NP in the complement of *believe* is indeed receiving Case as if it were the object of *believe* is supported by various diagnostics, for example, the loss of accusative in this position when *believe* is passivized, and the somewhat murky adjacency requirement between *believe* and the NP that is characteristic of objective case assignment in English (see below). The lexical subject of the non-finite complement of *believe* avoids the fate of other infinitival subjects (i.e. limitation to PRO) because in this configuration (which would come to be known as *Exceptional Case Marking* or 'ECM'), it is subject to Case assignment.

In sum, Case Theory constituted a significant advance in deducing the major effects of a rather puzzling, essentially descriptive, filter (the **NP-to-VP* filter),

[3] This same property (being Case assigners) accounts for why, in English, only V and P may take NP complements. Adjectives and nouns are not case assigners, and thus are limited to PP and CP complements, requiring *of* where a corresponding verb might take an NP complement (cf. *refuse the offer* vs. *refusal of the offer*). For an alternative approach to complementation asymmetries, within the Minimalist Program but without appeal to Case, see Pesetsky and Torrego 2004.

[4] Proposals in the GB framework for allowing Case assignment across the infinitival clause boundary vary in details, but share the common core that the complement of the *believe* class verbs is (or becomes, after a deletion operation) smaller than the complement of the *decide* class (control verbs). Updating the notation, the proposal in Chomsky (1981) is that infinitival complements are CPs, but that *believe*-type verbs induce a rule of CP-deletion, hence taking IP-complements at the level relevant for the application of the Case Filter. In later work, it is suggested that *believe* selects an IP rather than a CP complement as a lexical property (Chomsky and Lasnik 1995: 112).

978–0–19–920647–6 04-Spencer-ch03 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 47 of 58 May 16, 2008 18:57

largely from independently motivated elements of the theory (as in (4)) together with a very broad, and not construction-specific assumption, the Case Filter (3).[5]

3.1.2 Extensions: Case Theory

The main interest in Case Theory in GB and on into Minimalism lies not in the original empirical result, but in the consideration of a variety of intricately connected consequences. The postulation of the Case Filter had ramifications well beyond the distribution of infinitival subjects.

For example, Case could now be seen as one of the driving forces of movement for a variety of constructions. Thus, a unified account of promotion to subject in passive (5), raising (6), and unaccusatives (7), became possible: in each construction, the NP in its original position is not governed by a case assigner, and thus in each configuration, the NP must raise to finite subject position in order to satisfy the Case Filter. The Case Filter thus becomes one of the answers to the perennial question of why movement occurs, in the examples at hand, with reference to why movement to subject position occurs in English.

(5) a. Lina was kissed *t* (by Leo).
b. Kai was believed [*t* to have won the soccer match.]
c. The birdcage was found [*t* empty].

(6) a. Lina seems [*t* to like her brother].
b. Lina is likely [*t* to fall asleep].

(7) a. Jeffrey's bus arrived *t*.
b. The tree fell *t*.

Note, though, that a Case-based account of movement in these configurations is largely redundant with another core postulate of GB, namely the *Extended Projection Principle* (Chomsky 1982: 10), which includes the requirement that every finite clause have a subject (compare the *Final-1 Law* in Relational Grammar; see Perlmutter and Postal 1983a). The question of whether this redundancy can be eliminated has long stood as a research problem in GB/Minimalism.[6]

[5] Case Theory provides a deduction of the major effects of the **NP-to-VP* filter, but does not have exactly the same empirical coverage. Some effects not subsumed under Case Theory, as well as some empirical problems for the **NP-to-VP* filter, largely having to do with infinitival relatives, are addressed independently in Chomsky (1980), where a more comprehensive analysis of infinitival relatives is proposed.

[6] Fukui and Speas (1986) propose to reduce the EPP to Case Theory, by assuming that nominative case assignment must be established in overt syntax, by movement (rather than government or Agree). This idea (sometimes called the Inverse Case Filter) has been advocated within the Minimalist Program, see Bošković (2002) and especially Epstein and Seely (2006). It is not clear that this involves a genuine reduction, inasmuch as the empirical scope of the EPP and classic Case Theory are not identical. The core of Case Theory is held to be universal, while the obligatory NP-movement to

978–0–19–920647–6 04-Spencer-ch03 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 48 of 58 May 16, 2008 18:57

Another early result of Case Theory, set out by Stowell (1981), regards the 'order of complements' problem. As shown in (8), for verbs in English that select for multiple complements, it is generally held that the NP argument must precede all other (PP, CP) arguments, at least in 'neutral' clauses (i.e. clauses that are not derived by, for instance, *Heavy NP Shift* or similar operations which are typically associated with a 'special' intonation).

(8) a. Maggie donated [$_{NP}$ her allowance] [$_{PP}$ to the charity].
b. *Maggie donated [$_{PP}$ to the charity] [$_{NP}$ her allowance].

There is no such ordering effect in the corresponding nominalizations, as shown in (9), suggesting that the restriction is syntactic, rather than semantic, in nature.

(9) a. [Maggie's donation [$_{PP}$ of her allowance] [$_{PP}$ to charity]] was nice.
b. ?[Maggie's donation [$_{PP}$ to charity] [$_{PP}$ of her allowance]] was nice.

Similarly, multiple PP complements to a verb may be fairly freely reordered with respect to one another, in contrast to (8), suggesting that the restriction is specifically about NP complements.

(10) a. Julia talked [$_{PP}$ to Jillian's father] [$_{PP}$ about Tommy].
b. ?Julia talked [$_{PP}$ about Tommy] [$_{NP}$ to Jillian's father].

As Stowell argues, the particular requirement that NP complements precede all others (in English) can be seen as a special case of the general requirement of Case Adjacency, whereby the direct object of an accusative Case assigner (V or P) must be adjacent to its assigner (abstracting away from parentheticals and the like).[7] Deriving the order of complements from Case Theory meant that these ordering restrictions could be eliminated from the phrase-structure rules, an important step in the move towards a generalized X′-Theory, now standard in GB/Minimalism.

The above paragraphs illustrate the way in which Case Theory, originally suggested as an alternative to some very construction-specific filters regarding infinitives, could be neatly applied to a wide range of phenomena, allowing for increased generality of the rules at each step.

subject position seen in English is a matter of parametric variation (cf. Chomsky 1981: 27–8, and for more detailed recent arguments that the EPP is parameterized within Germanic, see Wurmbrand 2006 and references therein, especially Haider 1993, 2006). An opposing direction for the elimination of the redundancy, represented within GB/Minimalism by Marantz (1991) and others subsequently, is that the EPP be retained as a language-particular property, and Case Theory be abandoned. See the discussion in section 3.2, below.

[7] Stowell also considers the strong tendency for finite sentential complements (CPs) to extrapose, for which he introduces the Case Resistance Principle, and the assumption that CPs may not bear Case.

3.1.3 Government

Although (4) suffices as a good first pass at a description of Case assignment rules, the articulation of the details proved to be an important area of inquiry, largely because grammatical functions (subject, object) have been taken in GB and related frameworks to be derivative, structurally defined notions (see Chomsky 1965). From this perspective, an adequate characterization of the domain of objective case assignment seemed to require reference to two notions, namely c-command and adjacency. In formalizing the domain of Case assignment to capture the results discussed above, Chomsky (1980: 25) defines the notion of *government*. The original formulation is given here:

(11) α is *governed* by β if α is c-commanded by β and no major category or major category boundary appears between α and β.

The proper definition of government, and in particular, the formulation of what it means for a category to 'appear' (or later to 'intervene') between α and β was one of the most significant technical questions of the 1980s (see Aoun and Sportiche 1983, Chomsky 1981, 1986, Lasnik and Saito 1984, 1992 and Rizzi 1990 for a sample of important works). In Chomsky (1980), the definition was intended to subsume linear adjacency as well as structural intervention, though in most later work the adjacency condition on government, and hence on Case assignment, was held to be derivable from other properties of the theory (see, among others, Johnson 1991). Within the GB framework, the government relation, motivated originally for Case, was proposed to be a key notion at work in some way or another in a variety of modules of the theory; a derivative notion, *proper government*, playing a central role in subject–object asymmetries in extraction and related topics for example (see Chomsky 1986, and work cited there, especially Huang 1982). Stepping back from the details of a rather extensive literature, there is nevertheless a sense in which Case assignment remained at the core of the notion of government, providing the observable instances of the phenomenon from which the operational notion is to be generalized. If an NP in some configuration is clearly dependent on a particular element (V or P) for case, then that configuration must constitute a configuration of government, and that can be used heuristically to set parameters on the bounds of the definition.

3.1.4 Summary: Case Theory in GB

In sum, Case Theory, comprising at its core the Case Filter and a description of Case assignment, such as (4), stood very close to the core of the GB framework, with tangible results that moved the theory forward, and new questions that it opened up. Ultimately the results of Case Theory would be dependent on successful articulation of the theory of Case assignment. The explorations of locality, couched

978–0–19–920647–6 04-Spencer-ch03 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 50 of 58 May 16, 2008 18:57

in terms of *government*, constituted one aspect of the careful formalization of (4). Another area of exploration considered the implications of cross-linguistic variation in Case assignment rules (see below).

Although the notion *government* and the various attendant notions provided a greater technical precision, the basic rules of Case assignment in (4) left open a variety of questions, which would later serve as the impetus for a body of work at the transition from GB to Minimalism. Among the topics which received significant attention were asymmetries between nominative and accusative (or other) cases, including: (i) accusative was assigned by a lexical head (V or P), but nominative was assigned by a functional head (finite Infl); (ii) accusative was assigned under c-command (head-complement), where nominative was assigned under m-command (head-specifier); and (iii) accusative assignment was subject, in English, to an adjacency condition (as noted above), while nominative assignment was not (cf. *They probably won't win*, where the auxiliary is assumed to be in Infl). In addition, an unresolved question was why finiteness should matter for the assignment of nominative case by Infl.

3.2 How abstract is abstract case?

Given the explicit nature of the emerging theory, the ramifications of variation in Case assignment properties (as discussed in detail throughout this volume) would have predictable consequences for observable phenomena, and to some degree this was exploited with successful results. For example, Chomsky (1981: 122–3), discussing the work of Burzio (1981), suggests that Italian has a rule of nominative Case assignment to the post-verbal position, where English lacks such a rule. Under a Case-Theoretic approach to movement in passive, as sketched above, this difference in Case entails that Italian lacks obligatory NP-movement to the preverbal subject position in the counterparts to the English examples in (5).

On the other hand, the foundations of GB Case Theory were taken head on in a landmark paper on Icelandic (Zaenen, Maling, and Thráinsson, 1985, henceforth ZMT). That paper demonstrated that the original solution to the puzzle in (1) was inadequate, and that the traditional notion of Case, even allowing for an 'abstract' Case (in the sense of case lacking morphological realization), was not the driving force in the distribution of NPs, neither in the control/ECM alternation nor in the promotion to subject position in passive, raising, and unaccusatives. Space limitations do not permit a thorough review of this work here, but the key point can be made with reference to a few examples.

Icelandic is a language with overt morphological case distinctions. Transitive constructions involving a nominative–accusative array work exactly like their English counterparts – i.e. all the key distributional properties that are held to be the purview of Case Theory (including ECM, the ban on lexical subjects of infinitives, etc.) are robustly attested in Icelandic as well. In contrast to English, however, Icelandic also has what has become known as *quirky case* subjects, that is, subjects which are marked with a case other than nominative. One such example is given in (12a). The verb *hjálpa* 'help' governs lexical dative case on the object (cf. (12b)) and this case is hence retained in the passive. Importantly, the dative in (12a) must be seen as the subject of the sentence, a fact for which ZMT provide ample and convincing evidence, building on earlier work such as Andrews (1976) and Thráinsson (1979). (On this score, Icelandic contrasts minimally with German, which has superficially similar examples, but in which the datives are not subjects.) If it is assumed that 'subjecthood' is a structural property associated with a particular syntactic position (e.g. the specifier of IP), one must conclude that the dative argument moves to this position. This, then, however, is in conflict with a theory which aims at deriving movement to subject position from the Case Filter. Since the dative arguments receive (lexical) case, they should automatically satisfy the Case Filter in (3), and hence there would be no need for them to move to receive Case.

(12) a. *Þeim / honum var hjálpað.*
them / him.DAT was.SG helped
'They/He were/was helped.' (ZMT 99, 96)
b. *Ég hjálpaði honum.*
I.NOM helped him.DAT
'I helped them.' (ZMT 98)

The point can also be made with the control versus ECM distinction, the distributional puzzle at the historical core of Case Theory. In Icelandic, as in English, the subject of an infinitive must not be an overt NP, except when the infinitival clause is the complement of an ECM verb, such as *telja* 'believe'. However, this distinction cross-cuts case distinctions. A quirky subject shows exactly the same alternation between PRO and lexical NP as other subjects, but the quirky subject clearly does not depend on *believe* for its (overt) case – the dative on the NP *þeim* 'them' in (13b) is assigned by the verb 'help', not by *tel* 'believe' (which assigns accusative).

(13) a. *Ég vonast til [að PRO verða hjálpað].*
I.NOM hope for to be helped
'I hoped to be helped.' (ZMT 109)
b. *Ég tel þeim hafa verið hjálpað í prófinu*
I.NOM believe them.DAT to.have been helped in exam.the
'I believe them to have been helped on the exam.' (ZMT 107)

The work of ZMT shows that Case, for the purposes of the Case Filter, cannot be equated with the morphological case that is realized on NPs (even allowing for

978–0–19–920647–6 04-Spencer-ch03 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 52 of 58 May 16, 2008 18:57

zero-realization, as on English nouns). ZMT thus argued for the abandonment of Case Theory (a direction pursued in the GB/MP framework by Marantz 1991 and others), and suggested that the relevant licensing theory be formulated directly in terms of GFs (they thus took their paper as support for LFG). Others within the GB framework (Cowper 1988, Freidin and Sprouse 1991) took the ZMT results as instead showing that the Case Filter was correct, but referred to an even more abstract notion of Case than envisioned in Chomsky (1980, 1981), one which may align with the morphologically visible case system in the basic patterns, but need not do so in all instances. Thus in Icelandic, 'abstract nominative' (the Case of finite subjects) could be realized as morphological nominative (the basic realization), accusative, dative, or genitive, and similarly, 'abstract accusative' (the Case of objects, including ECM objects) could also be realized by any of the four morphological cases. Largely missing from the debate, so far as we know, are concerted attempts to show that these two perspectives are distinct in substance, and that the postulation of 'abstract nominative', for example, is distinct from the postulation of 'subject' as a grammatical primitive. Steps towards untangling these views have appeared only quite recently (see in particular Legate 2005, and for an opposing view Bobaljik, to appear).

3.3 Case Theory in the Minimalist Program

The early 1990s marked a transition from GB to Minimalism as the declared framework of choice for many generative syntacticians. A good deal of confusion seems to have arisen around nomenclature. Throughout the Minimalist writings, Chomsky has been at pains to stress that Minimalism is a *program*, not a theory or theoretical framework. That is, in contrast to formal approaches to this point, which have focused on providing explicit characterizations of grammar (universal and particular) at increasing levels of abstraction and generalization, Minimalism has been proposed as a set of rough guidelines about the types of questions that may advance linguistic inquiry into the domain of asking why UG is the way it is. Despite the new focus underlying the Minimalist Program, the sizeable majority of work cast as Minimalist is characterized by a renewed focus on uniformity and theoretical parsimony, but the goal on the horizon in most of this work is still (as in GB) an explicit specification of the principles and parameters of UG, in other words, the limits of what it means to be a 'possible language'. We turn now to a brief outline of where Case Theory fits within Minimalism in practice, the next phase of the Principles and Parameters framework, and then examine where Case

978–0–19–920647–6 04-Spencer-ch03 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 53 of 58 May 16, 2008 18:57

Theory fits in the (currently programmatic) discussion of what is known as the *Strong Minimalist Thesis*.

3.3.1 Minimalism in practice

For Minimalism in practice, the major questions of Case Theory revolve around the differences between nominative and accusative case assignment identified above (section 3.1.4), and thus, the possibility of developing a uniform theory of nominative and accusative case assignment.

This has proven to be a productive domain of inquiry, with a variety of proposals for a unified theory of Case assignment on offer. One perspective (exemplified prominently by Chomsky 1991) proposes to assimilate accusative Case assignment to the same type of structural configuration as nominative, namely m-command, or in more current terms a spec(ifier)-head relation (see Koopman 2006 for a recent defense, and Wurmbrand 2006 for empirical problems in generalizing spec-head even in the limited domain of nominative subjects in Germanic). At the other end stands the proposal that all case assignment, including nominative to the subject, should be characterized by a government relation, that is, c-command and locality (see Chomsky 2000, where the relationship is termed Agree). The door to this possibility is opened by assuming that all subjects in Spec,IP are moved there from lower positions, a proposal originally suggested by Ken Hale (class lectures) and Koopman and Sportiche (1991). Note that the Agree perspective, while unifying the mechanisms involved in structural Case assignment, shifts the burden of the motivation for movement in examples like (5)–(7) away from Case Theory, to the EPP or an extension thereof.

An additional line of investigation on the theme of a uniform case assigning (or 'case-checking', a distinction we set aside here) mechanism concerns the functional versus lexical difference in the case assigners Mahajan (1989), Déprez (1989), Chomsky (1991), Johnson (1991), and others postulate VP-external functional projections responsible for Case on objects, leading to a uniform proposal that Case is assigned by functional heads. One striking piece of empirical evidence for this comes from Long Passive (and Long Unaccusative) in German, Spanish, Japanese, and other languages. Relevant examples are given in (14) and (15) – embedded clauses are used to avoid the additional complications posed by Verb Second. The key point to observe is that it is the voice distinctions on the embedding verb (in this case *versuchen* 'try') that determine the possibility of a nominative/accusative case alternation on the object of the embedded verb. The alternation is passive-like, but it is the voice marking associated with the higher verb only that determines the case properties in the lower domain. See Wurmbrand (2001) for extended discussion and additional references.

(14) *weil er den/*der Traktor versucht hat [t_{OBJ} zu reparieren]*
since he the.ACC/*NOM tractor tried has [t_{OBJ} to repair]
'since he tried to repair the tractor'

(15) a. *weil der Traktor zu reparieren versucht wurde*
since the.NOM tractor to repair tried was
'since they tried to repair the tractor'
b. *weil die Traktoren zu reparieren versucht wurden*
since the tractors (NOM) to repair tried were
'since they tried to repair the tractors'

Though there has been a veritable explosion of proposals, it seems clear that some reduction of the apparent differences between nominative and accusative mentioned in section 3.1.4 is a real possibility, within the GB/Minimalist frameworks.

3.3.2 Programmatic Minimalism

In addition to the technical innovations that characterize Minimalism in practice, a small minority of Minimalist work poses questions that go beyond specification of the principles and parameters of UG, and toy with the key Minimalist question: how close does language come to 'optimal design', where 'optimal design' is to be understood as having no properties other than those dictated by the need for the language module to interface with two other cognitive systems: the conceptual-intensional system and the motor-articulatory system. The *Strongest Minimalist Thesis* (SMT, Chomsky 2001: 1) is the conjecture that language approaches optimality in this very particular sense. Case, in particular Abstract Case, should leap to the foreground in such a line of inquiry, in the form of the question in (16).

(16) Why should there be Abstract Case at all?

At first blush, Abstract Case seems problematic for the *SMT*, in the sense that there is no obvious interface pressure for its existence. To the extent that there is any temptation to toy with functionalist explanations, such as the *identifying* and *distinguishing* functions of case recognized by Mallinson and Blake (1981), Comrie (1989) and others, these seem ill-suited to the core examples of Abstract Case at work in languages such as English where Case is not marked.

One speculation (see e.g. Chomsky 2004) is that abstract Case features, now generalized as part of a system of uninterpretable formal features, lie at the heart of the linguistic coding of what Chomsky refers to as the 'duality of semantics' (Chomsky 2004: 7), one side being thematic relations/argument structure, and the other being information structure and scopal relations. Under the most recent conception, Case features allow for the proper working of the Probe-Goal system, a feature-checking mechanism that is in a loose sense the descendant of licensing under government.

Another speculation (see Pesetsky and Torrego 2001) is that what we call Case is in fact the NP analogue of tense in the verbal system. As pointed out above, at the time of writing, that aspect of the Minimalist Program which focuses on the SMT and questions at a similar level is, according to Chomsky, a research program, still in its infancy. We believe it is fair to say that the jury is still out as to whether this perspective will yield new insights in this unexplored terrain, in large part due to a paucity of evidence concerning the properties of the interfaces.

3.4 Topics in Case Theory within GB/Minimalism

In closing, we note briefly three additional topics that have received prominent attention from the perspective of Case Theory within GB and/or Minimalism. In some sense, each of these topics constitutes an add-on to the core Case Theory, in that the proposals cited have had less influence on the direction Case Theory has taken than the topics considered above.

3.4.1 Null Case

The ECM/control distinction (1a–b) at the historical heart of Case Theory received renewed attention in the early Minimalist period. The original analysis relied on three stipulations: (i) a selectional difference for the infinitival complements of ECM (IP) vs. control (CP) verbs; (ii) the absence of a Case assignment rule to the subject position of infinitives (see (4)); and (iii) exempting the NP PRO from the Case Filter. One strand of inquiry asked whether these stipulations might not be derivable, at least in part. The major investigation in this area centred on Stowell's (1982) observation (developed in more detail in Pesetsky 1992) that control infinitives are typically future irrealis infinitives, whereas ECM infinitives are typically propositional infinitives.[8] Building on this observation, Chomsky and Lasnik (1995), Bošković (1996, 1997), and Martin (1996, 2001) proposed an amendment to (4) whereby the subjects of 'tensed' infinitives (those with a future irrealis interpretation) do assign Case to their subject position, but the Case assigned is a special 'Null Case' that only PRO (but not lexical NPs) may bear. Under this view, the Case Filter could thus be taken to regulate all NPs including PRO, an (apparent) reduction in the stipulations needed to account for the difference. The Null Case

[8] As shown in Pesetsky (1992), the picture is, in fact, far more complex. However, we abstract away from various complications here.

978–0–19–920647–6 04-Spencer-ch03 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 56 of 58 May 16, 2008 18:57

approach has proven controversial within the Minimalist Program, though (see Landau 2000, Baltin and Barrett 2002, Hornstein 2003, Cecchetto 2004, Wurmbrand 2005, To appear, for a variety of empirical and theoretical challenges). One of the most serious problems raised for the Null Case view is that the presence vs. absence of 'infinitival tense' – the predictor of Null Case – has itself not been shown to be predictable on independent grounds, leaving Null Case as possibly simply a notational variant of the earlier account.

Another challenge for Case Theoretic treatments of the ECM/Control distinction, (problematic for both the Null Case view and the original Case Theoretic analysis in section 3.1) comes from languages in which the subject of (non-ECM) infinitives receives a detectable morphological case (not Null Case). Once again, Icelandic has played an important role in the discussion since morphological case on PRO is detectable via elements that show agreement in case with the subject position (see Sigurðsson 1991). A different aspect of this problem (noted already in Chomsky 1981: 140, n. 25) is posed by languages that allow overt, case-marked subjects of non-finite, non-ECM clauses, such as accusative subjects in Latin and Greek.

3.4.2 Ergativity and case typology

Within GB (and to a lesser degree in Minimalism), Case Theory was dependent upon rules of Case assignment, which were held to be a point of cross-linguistic variation (see section 3.2). Perhaps the most striking aspect of cross-linguistic variation in (morphological) case lies in the existence of alignments other than nominative–accusative, including ergative (and split-ergative) systems (Bickel and Nichols, this volume). Ergativity received a treatment early in the GB period in Marantz (1981), with a variety of subsequent proposals in both the GB and Minimalism frameworks (see Butt 2006: chapter 6 for a survey and Johns, Massam, and Ndayiragije 2006 for a collection of recent views).

Perhaps the most important question of this literature (still unresolved) is whether ergativity is a syntactic or morphological phenomenon, in other words, whether the different morphological case patterns are correlated with any syntactic differences. This was a particularly important question in the early GB period, when (structural) Case was equated with (morphological) case. In that framework, since the rules of case assignment in an ergative system are different from those in a nominative–accusative system, this morphological difference should have repercussions, via Case Theory, for the syntax. For example, in an ergative system (by definition) the case assigned to an intransitive subject is the same as that assigned to the object (not the subject) of a transitive clause. Within GB and Minimalism, subject and object cases are distinct in many syntactic ways. One such difference is that the subject case, but not the object case, was held to be tied to the finiteness

of the clause (but see Sigurðsson 1991). In a Case-Theoretic approach to ergativity, this difference might be expected to play out as a restriction whereby only transitive subjects are obligatorily suppressed in infinitives, but where intransitive subjects pattern with objects in being freely expressed, even in infinitives. Bobaljik (1993a), developing ideas from Levin and Massam (1984), explores a pattern of agreement in Inuit languages that may be described in these terms. The relevance of ergativity to Case Theory becomes less obvious, though, with the post-ZMT recognition of a more abstract relation between Structural Case (the abstract, syntactic licensing) and morphological case. Like quirky case, this further level of abstraction leaves open the possibility that ergativity is best described as a morphological phenomenon, lying squarely outside the domain of Case Theory, and masking a (more) uniform syntax. We note as well that more extensive investigation of a variety of languages with ergative case and/or agreement systems suggests that ergativity is not a uniform phenomenon, and that there is considerable syntactic variation among languages with ergative case systems, and hence the questions just raised must be asked not for ergativity as a phenomenon, but for individual languages or groups thereof.

3.4.3 Case, agreement, and beyond

A final important topic that has been explored within GB and Minimalism concerns the relationship of Case and case to other features of the grammatical system. The position taken by Chomsky in the Minimalist writings is that case and agreement are instantiations of the same fundamental grammatical relationship, representing head versus dependent marking of that relation (although what the precise relationship is has shifted over the course of the Minimalist period).[9] Here, as above, Icelandic evidence has played a prominent role, since in Icelandic agreement is indeed intricately connected to (morphological) case: only nominative NPs may govern agreement on the finite predicate, and they do so regardless of grammatical function: nominative objects trigger agreement, while non-nominative subjects do not (see Sigurðsson 1996; for evidence that it is case and not grammatical function that determines agreement controller in a variety of languages, see Falk 1997 and Bobaljik, to appear). The Icelandic evidence is somewhat of a two-edged sword, though, since it is precisely the non-nominative subjects that have the distribution

[9] Inasmuch as the relationship between (structural) Case and (morphological) case is transparent, this view would be simply incorrect in light of the many well-attested 'mismatches' between case and agreement Such mismatches include the type of ergative split in Warlpiri or Chukchi in which ergative–absolutive case marking occurs alongside subject–object (nominative–accusative) alignments in agreement (see Dixon 1994), and also more intricate mismatches such as Basque Ergative Displacement (Laka 1993, Hualde and Ortiz de Urbina 2003), and the Chukchi Spurious Antipassive (Spencer 2000, Hale 2002, Bobaljik and Branigan 2006). See Legate (2005) for one attempt to defend the canonical Minimalist view in light of such apparent counter-evidence.

978–0–19–920647–6 04-Spencer-ch03 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 58 of 58 May 16, 2008 18:57

attributed, in GB, to abstract nominative structural Case, while the nominative objects were (implicitly, at least) treated as bearing abstract accusative. For recent perspectives on nominative objects and review of the current literature, see Hiraiwa (2005) and Nomura (2005).[10]

The trajectory exemplified by the focus on the abstract features underlying case and agreement is arguably carried a step further in Chomsky's later Minimalist writings. For example, Chomsky (2000) signals a shift in emphasis, if not in substance, to a more broad category of 'uninterpretable features'. In many ways, the latest position advocated by Chomsky has come ever closer to the position forwarded by Marantz (1991), a position which maintains, from GB's Case Theory/Case Filter, a component of formal licensing for NPs, not deducible from semantic or phonological (i.e. 'interface' requirements), but for which case plays little role, beyond the observation that case and agreement are among the possible morphological signals of this formal licensing. We close with a passage from Chomsky (2000) setting out this view.[11]

> According to this conception, agreement (hence movement) is driven by uninterpretable features of the probe, which must be deleted for legibility . . . With this shift in perspective, structural Case is demoted in significance. The Case Filter still functions indirectly in the manner of Vergnaud's original proposal, to determine the distribution of noun phrases. But what matters primarily are the probes, including φ-features [person, number, gender – B&W] of T, v. That reverses much of the recent history of inquiry into these topics and also brings out more clearly the question of why Case exists at all. The question arises still more sharply if matching is just identity, so that Case can never be attracted; operations are not induced by Case-checking requirements. (Chomsky 2000: 127)

[10] As in many domains, subtle differences in the use of key terms (abstract, inherent, lexical, structural, and quirky) among authors have contributed to the complexity of the debate.

[11] Although we focus on Chomsky's writings as the canonical instantiation of Minimalism, it is well worth noting, as Chomsky does, that there are a variety of current, mutually incompatible, theoretical proposals in the Minimalist literature that share certain core assumptions. With regards to Case Theory, Legate (2005) and Bošković (to appear) maintain a view of Case more in line with later GB proposals than suggested in the passage from Chomsky.

CHAPTER 4

CASE IN LEXICAL-FUNCTIONAL GRAMMAR

MIRIAM BUTT

4.1 Introduction

The treatment of case has been one of the central concerns within Lexical-Functional Grammar (LFG) since its inception in the late 1970s. Several of the papers collected in the seminal book marking the emergence of LFG (Bresnan 1982) deal with case marking in particular (e.g., Neidle on Russian, K. P. Mohanan with respect to Malayalam). However, a sophisticated analysis of cross-linguistic case patterns only became available as LFG's Linking Theory (known as Lexical Mapping Theory) evolved. In particular, once it was recognized on the basis of argumentation by Rappaport (1983) that argument structure needed to be posited as a level of representation that was independent of constituent structure (e.g. very much unlike the assumptions of GB/MP, see Chapter 3), the way was paved for analyses of case to be stated in terms of generalizations over a(rgument)-structure. That is, in terms of generalizations that take both *semantic* and *syntactic* factors into account.

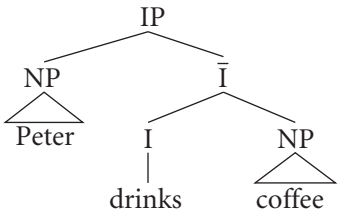
This chapter first presents some LFG basics in section 4.2, then briefly charts the development of Linking Theory in section 4.3 and presents current theories of case in section 4.4. The chapter closes with a consideration of LFG-based analyses that are couched within Optimality Theory (see Chapter 6 of this volume) in section 4.5.

4.2 LFG BASICS

LFG separates facts about linear word order and constituency from the functional dependency structure analysis of a clause. Word order and constituency are represented at c(onstituent)-structure via tree representations, as shown in (1). LFG assumes a version of X′-theory that goes back to Bresnan (1977). For current assumptions about c-structural representations, see Bresnan (2001).

The f(unctional)-structure is represented in terms of an attribute-value matrix (AVM) and encodes functional syntactic information about grammatical relations, tense/aspect, case, number, person, etc. A sample f-structure for (1) is shown in (2).

(1) Peter drinks coffee.



(2) Peter drinks coffee.

$$\begin{bmatrix} \text{PRED} & \text{'drink<SUBJ,OBJ>'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'peter'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'coffee'} \\ \text{PERS} & 3 \\ \text{CASE} & \text{ACC} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \end{bmatrix}$$

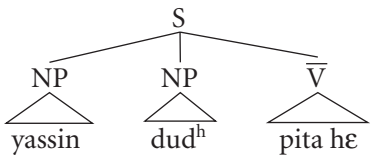
F-structures are a *projection* from the c-structure because they are related to the c-structure via a formal system of annotations. The effect of the projection architecture is that the levels of representation *constrain* each other mutually. That is, an analysis can only be successful if the f-structure information is complete and consistent, and if the phrase structure rules license the structure. Information at f-structure may flow together from different sources (not illustrated here). The pieces of information are combined with one another via *unification*. In contrast to the fundamental derivational assumptions of GB/MP (Chapter 3 of this volume), LFG assumes no derivations from one structure to another. Indeed, this is one of the characteristics which makes LFG computationally tractable.

The f-structure is thus an abstract level of representation which is not tied to the particular word order or surface form of a language. Cross-linguistic generalizations about passivization, for example, are formulated with respect to

f-structure. The c-structures, in contrast, encode language-specific requirements on word order and constituency. The idea is that an SVO language like English and an SOV language like Urdu may differ wildly on the surface, but are similar at the basic predicational level, as shown in (4) and (5). The c-structure analysis in (3) differs from the English one in (2) because Urdu is a free word order language in which the NPs of a sentence are sisters (there is no evidence for a VP).[1]

(3) *yassin* *dud^h* *pi-t-a* *hɛ*
Yassin.M.Sg.Nom milk.F.Sg.Nom drink-Impf-M.Sg be.Pres.3.Sg
'Yassin drinks milk.' Urdu

(4)



(5)

$$\begin{bmatrix} \text{PRED} & \text{'pina<SUBJ,OBJ>'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'Yassin'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 3 \\ \text{GEND} & \text{MASC} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'dud}^h\text{'} \\ \text{GEND} & \text{FEM} \\ \text{PERS} & 3 \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \end{bmatrix}$$

4.2.1 Grammatical functions

The above f-structures include two basic grammatical relations: SUBJ(ect) and OBJ(ect). Grammatical relations are assumed as part of the syntactic inventory of every language and are referred to as *grammatical functions* (GF) to indicate their functional status, which is the relation of arguments and predicational elements to one another. LFG assumes the GFs in (6). Like many other elements in linguistics, GFs can be arranged in a hierarchy, whereby the SUBJ is the 'highest' on the hierarchy, the OBJ the next highest, etc.

[1] Within LFG, only minimal constituents are assumed, for which there is actual distributional evidence in the language. The difference between a language like English and a language like Urdu is taken to be a difference between endocentricity (all heads project maximally) and exocentricity (heads do not necessarily maximally project). See Bresnan (2001, Ch. 7) and Dalrymple (2001, Ch. 3) for further discussion.

(6) **Grammatical Functions**
SUBJ OBJ OBJ_θ $\text{OBL}(\text{ique})_\theta$ COMP(lement) XCOMP(lement) ADJUNCT

Dalrymple (2001: 11–27) provides a useful discussion of the GFs as well as several syntactic tests by which they can be identified. GFs are not inherently identified with any particular case in LFG. That is, there is no assumption that SUBJ will always be nominative and that OBJ will always be accusative. This is because of the early recognition of non-nominative subjects due to Zaenen, Maling, and Thráinsson's (1985) LFG-based investigation of Icelandic.

Of course, individual languages may prove to have strict ideas about the association of case and GFs. For example, it is claimed for German that subjects are always nominative and that nominatives are always subjects. These correlations are taken into account on a language-particular basis within LFG, but are not assumed to necessarily be cross-linguistic universals.

4.2.2 Argument structure and thematic roles

In addition to the basic c- and f-structural representations, LFG's projection architecture potentially allows for several other projections. One standard additional projection is the s(emantic)-projection (e.g. Halvorsen and Kaplan 1988, Dalrymple 1999), which encodes the semantic analysis of the clause. The a-structure can also be conceived of as a projection, it can also be seen as an elaboration of the PRED value in f-structures such as (2) or (5) (Alsina 1996). A-structure in LFG in theoretical papers is generally represented as in (7), though it can also be formally represented as an AVM (Butt 1998), in line with f-structure representations.

(7) *pinch* < agent theme >

The a-structure encodes predicate–argument relationships in terms of thematic roles. These thematic roles are generally arranged in a thematic hierarchy, shown in (8) (based on Bresnan and Kanerva 1989). However, in practice very little reference is ever made to the hierarchy, except to the agent as being highest.

(8) **Thematic Role Hierarchy**
agent > *beneficiary* > *recipient/experiencer* > *instrumental* > *theme/patient* > *location*

4.3 Linking Theory

The first paper to associate case marking patterns with a-structure information was also the first paper to formulate some initial linking principles. Zaenen, Maling,

and Thráinsson (1985) looked at case marking patterns in Icelandic. Besides devising tests that conclusively established the existence of non-nominative subjects, Zaenen, Maling, and Thráinsson (ZMT) formulated principles which governed a complex relationship between thematic roles, case, and grammatical functions. As an example, their association principles for Icelandic are shown in (9).

(9) **Icelandic Association Principles**
1. AGENTS are linked to SUBJ (Universal)
2. Casemarked THEMES are assigned to the lowest available GF. (Language-specific)
3. If there is only one thematic role, it is assigned to SUBJ; if there are two, they are assigned to SUBJ and OBJ; if there are three, they are assigned to SUBJ, OBJ, 2OBJ [NB: the modern OBJ_θ]. This principle applies after principle 2 and after the assignment of restricted GFS. (Universal)
4. Default Case-Marking: the highest available GF is assigned NOM case, the next highest ACC. (Universal)

The effect of the association principles is best illustrated via a concrete example. The Icelandic verb *óska* 'to wish' can be used either transitively or as a ditransitive so that the goal 'her' in (10) is optional. When it is present, it is realized as the direct object (OBJ). When it is not present, the theme argument is instead linked to the direct object, as shown in (11).

(10) *Þú hefur óskað (henni) þess*
you have wished her.DAT this.GEN
'You have wished this on/for her.' (ZMT 1985: 470) Icelandic

(11)

óska: <	agent	theme [+gen]	(goal) > [+dat]
a.	SUBJ	2OBJ	OBJ
b.	SUBJ	OBJ	

ZMT define inherent lexical case as being an idiosyncratic property of a lexical item, assigned by a verb, preposition, or adjective. In (11), the theme 'this' is assigned inherent genitive case by the verb 'wish'. If the theme were not marked with the genitive feature, then it would be linked to an OBJ in both the transitive and the ditransitive scenario by the association principles in (9). But given the special marking in conjunction with the language-specific principle in (9.2), it is assigned to a secondary object rather than the direct object in (11a). In (11b), the theme is linked to the direct object because that is the lowest available GF, given that there are no further arguments to be accommodated.

ZMT's notion of inherent case came to be known as quirky case. The term 'quirky' suggests a random lawlessness, but a close inspection of ZMT's original paper shows that inherent case assignment actually proceeds in a very regulated manner. The 'quirky' genitive or dative cases are always regularly associated with a

given thematic role. Genitives regularly occur on themes (11) and datives mark goals as well as themes. There seem to be no instances of truly idiosyncratic case, rather case assignment seems to be principled and follows from lexical semantic factors.

Today's standard Linking Theory relates GFs to thematic roles via two abstract linking features, [±o](bjective) and [±r](estrictive), by which both thematic roles and GFs can be classified. Additionally, a number of principles govern the association of GFs and thematic roles. Furthermore, argument changing operations such as passives (argument deletion) or applicatives (argument addition) are taken into account (see Butt 2006 for a detailed discussion). The role of case in most accounts has stayed much as it was in ZMT's analysis of Icelandic: an extra piece of information that helps determine the mapping between GFs and thematic roles. However, there are also some approaches that take the theory of case within LFG a bit further. Two of these are sketched briefly in the next section.

4.4 Some theories of case

4.4.1 Constructive Case

Nordlinger (1998b, 2000) takes on the phenomenon of *case stacking* in Australian languages and develops a theory of *Constructive Case*. The phenomenon is exemplified by (12), where the word 'pouch' is marked with three cases: one to show that it is signalling a location, one to show that it is part of a possessive or accompanying relation to another word (the proprietive case), and one to show that it is part of (modifying) an accusative case-marked noun (see Chapter 52 for more discussion of this type of case marking). The word 'joey' (a baby euro – a type of kangaroo) has two cases. The proprietive shows that it stands in an accompanying relationship with another (with the euro), and the accusative to show that it is part of (modifying) an accusative case-marked noun. Finally, the 'euro' is accusative as the direct object, while the 'I' is nominative (unmarked).

(12) *Ngayu nhawu-lha ngurnu tharnta-a mirtily-marta-a*
I saw-PAST that.ACC euro-ACC joey-PROP-ACC
thara-ngka-marta-a.
pouch-LOC-PROP-ACC
'I saw the euro with a joey in (its) pouch.' Martuthunira
(Dench 1995: 60)

The f-structure analysis in (13) shows that the case markers indicate which functional layers of analysis the case-marked word is embedded in. That is, the three case markers on 'pouch' signal that it is a locative adjunct embedded under a proprietive adjunct that in turn modifies an accusative direct object.

Nordlinger (1998b) sees the case morphology itself as playing a large role in *constructing* the syntax of the clause. An ergative case in her analysis, for example, carries the following pieces of syntactic information: i) that there is a subject; ii) that it is ergative (ergatives are always subjects, but subjects are not always ergative). These pieces of information are encoded as part of the lexical entry of the ergative, as shown in (14a). Similarly, the abstract entries for the accusative, proprietive, and locative in Martuthunira specify the grammatical function they indicate, as well as the case feature, as shown in (14b–d).

Note the special use of the ↑ in the lexical entries of the case markers. The specification of the case feature is standard: each case marker specifies that the attribute CASE is assigned a certain value (ergative, accusative, etc.). This ensures that whatever constituent carries the case marker will be analysed as ergative, or accusative, or locative, etc. The second line in each entry involves *inside-out functional designation* (Dalrymple 1993, 2001). The ↑ following the specification of a grammatical function formulates a requirement that, come what may, the constituent should be analysed as a subject in (14a), an object in (14b), and an adjunct in (14c–d).[2]

(13)
$$\begin{bmatrix} \text{PRED} & \text{'see<SUBJ,OBJ>'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{NUM} & \text{SG} \\ \text{PERS} & 1 \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'euro'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{CASE} & \text{ACC} \\ \text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{'joey'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{CASE} & \text{PROP} \\ \text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{'pouch'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{CASE} & \text{LOC} \end{bmatrix} \right\} \end{bmatrix} \right\} \end{bmatrix} \\ \text{TENSE} & \text{PAST} \end{bmatrix}$$

[2] See also Andrews (1996).

(14) a. ERGATIVE: (↑CASE) = ERG
(SUBJ ↑)
b. ACCUSATIVE: (↑CASE) = ACC
(OBJ ↑)
c. LOCATIVE: (↑CASE) = LOC
(ADJUNCT ↑)
b. PROPRIETIVE: (↑CASE) = PROP
(ADJUNCT ↑)

Now consider the example in (15) from Wambaya. Here the NP 'big dog' is a discontinuous constituent. This poses no problem for Nordlinger's view of case as the effect of the analysis is that the combination of information from the lexical entries of 'big', 'dog', and the ergative case in (14a) results in the two partial f-structures shown in (16) and (17). Both the ergative 'dog' and the 'big' specify that they are parts of the subject. The 'dog' serves as the head of the phrase and the 'big' as an adjunct which modifies it.

(15) *galalarrinyi-ni* *gini-ng-a* *dawu* *bugayini-ni*
dog.I-ERG 3SG.MASC.A-1.O-NFUT bite big.I-ERG
'The big dog bit me.' Wambaya

(16) $\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'dog'} \\ \text{CASE} & \text{ERG} \end{bmatrix} \end{bmatrix}$

(17) $\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{ADJUNCT} & [\text{PRED} \ \text{'big'}] \end{bmatrix} \end{bmatrix}$

These two sets of information are unified into the structure shown in (18) as a routine part of the clausal analysis within the LFG formalism. The problem of discontinuous constituents is solved by using the case morphology as a primary source of information about clausal structure.

(18) $\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{ERG} \\ \text{PRED} & \text{'dog'} \\ \text{ADJUNCT} & [\text{PRED} \ \text{'big'}] \end{bmatrix} \end{bmatrix}$

However, note that Nordlinger's system makes no mention of a-structure and, in contrast to ZMT on Icelandic, draws no connection between a-structural information and case marking. The next section introduces a model of case within LFG that assumes a complex interaction between a-structure, GFs, and case marking.

4.4.2 Differential Case Theory

Urdu exhibits alternations in which the only difference between two clauses is the case morphology on one of the noun phrases. This morphological difference

signals a difference in semantic interpretation. Some examples are shown in (19) and (20). In (19a), the ergative subject indicates that the participant ('Nadya') has some control over the action, whereas the dative case in (19b) implies more of an obligation to perform the action (see Bashir 1999 for a more differentiated analysis). This is an example of case marking involving a *subject alternation*, or *differential subject marking* as it has come to be called in the Optimality Theoretic literature (see Chapter 6).

(19) a. *nadya=ne* *zu* *ja-na* *hɛ*
Nadya.F.SG=ERG ZOO.M.SG.OBL go-INF.M.SG be.PRES.3.SG
'Nadya wants to go to the zoo.' Urdu

b. *nadya=ko* *zu* *ja-na* *hɛ*
Nadya.F.SG=DAT ZOO.M.SG.OBL go-INF.M.SG be.PRES.3.SG
'Nadya has to go to the zoo.' Urdu

Example (20) involves an *object alternation*, or *differential object marking* in Optimality Theoretic terms. Here the only difference between the two sentences is the case marker on the object in (20b). In addition to marking dative case, as in (19b), the *ko* also functions as a marker of specificity/definiteness on direct objects in Urdu. Because it appears only on direct objects in this semantic context, and because it is not retained under passivization, this case marker can be analysed as an accusative. The effect of *ko* in (20b) is that Nadya must be interpreted as having a particular giraffe in mind that she wants to go to see. In (20a), in contrast, it could be some generic giraffe or giraffes that Nadya would like to see (at the zoo, for example). The 'giraffe' in (20a) is glossed as nominative. This case has no overt morphophonological realization in Urdu, something which is cross-linguistically quite common.

(20) a. *nadya=ne* *jiraf* *dek^h-na* *hɛ*
Nadya.F.SG=ERG giraffe.M.SG.NOM see-INF.M.SG be.PRES.3.SG
'Nadya wants to see a giraffe/giraffes.' Urdu

b. *nadya=ne* *jiraf=ko* *dek^h-na* *hɛ*
Nadya.F.SG=ERG giraffe.M.SG=ACC see-INF.M.SG be.PRES.3.SG
'Nadya wants to see the giraffe.' Urdu

Parallels to this alternation between nominative and accusative can be found in Turkish (Enç 1991) and other South Asian languages. In this example, as in (19), the only difference between the a and b versions is the case marking. This, and the clear connection to a semantic difference, prompted Butt and King (1991) to begin formulating an approach to case that included a notion of *semantic case*. This term has generally been used to refer to the case marking of adjuncts such as locatives or temporal expressions. Butt and King used this term to apply to those case markers of core arguments which also contribute information that is relevant for the final semantic interpretation of the clause. In order to allow the semantic information contributed by the case markers to flow directly into the analysis of the clause, Butt

and King (1991), exactly like Nordlinger (1998b), proposed explicit lexical entries for case markers. An example taken from later work (Butt and King 2003), is shown in (21) for the use of accusative *ko* in (20b).

(21) ko (↑ CASE) = ACC
(OBJ ↑)
(↑*sem-str* SPECIFICITY) = +

Butt and King's (1991) proposals for semantic case foreshadowed Nordlinger's ideas about *constructive case* in that the case markers themselves are considered to be active components which contribute to the analysis of a clause. The lexical entry for the accusative use of *ko* states that: i) the case is accusative; ii) the relevant NP should be a direct object; iii) the NP should be interpreted as specific at the level of s(emantic)-structure.

But the information carried by case markers is only part of a more complex system that accounts for differential case marking patterns. Butt and King (2003, 2005) assume the version of linking theory proposed in Butt (1998). In this version of linking theory, no explicit thematic hierarchy is assumed and the thematic roles are restricted to a very basic set: *agent*, *goal*, *theme*, *locative*. Beneficiaries, recipients, experiencers, etc. are all assumed to be an instance of a goal, more or less abstract.

As in standard linking theory, case as such is not integrated directly into the linking between GFs and a-structure. However, case marking becomes relevant for linking when different linking possibilities exist, just as was the case for ZMT. The information provided by the case morphology can be used to select just one of the possibilities.

But, as the examples in (19) and (20) show, the function of case marking must go beyond just the determination of grammatical functions. In fact, under Butt and King's Differential Case Theory (DCT), most instances of case work out to be instances of semantic case, which is defined as being simultaneously: i) semantically predictable; ii) subject to syntactic restrictions, such as being confined to certain grammatical functions. Indeed, the bulk of the cross-linguistic case marking phenomena involve an interaction between syntactic and semantic constraints.

Take the ergative case in Urdu, for example. This case marker can only appear on subjects and so must obey a particular syntactic restriction. Because it is an agentive case, the ergative does not just play a structural role, it also gives rise to some semantic effects. As is well known from a host of studies on Hindi/Urdu, the ergative alternates with the unmarked nominative on unergative intransitives (the most comprehensive study to date of the distribution of the ergative is Davison 1999). As shown in (22), this alternation correlates with an expression of control/volitionality.

(22) a. *yassin* *k^{h}ās-a*
Yassin.M.SG.NOM cough-PRF.M.SG
'Yassin coughed.' Urdu

b. *yassin=ne* *kʰãs-a*
Yassin.M.SG=ERG cough-PRF.M.SG
'Yassin coughed (purposefully).' Urdu

Another example is the accusative/dative *ko* in Urdu. As was seen in connection with example (20), when it appears on direct objects, it signals specificity. Again, there is a combination of syntactic (direct objects only) and semantic factors (specificity). The *ko* can also appear on subjects (19b) and on indirect objects, as in (23).

(23) *nadya=ne* *bɪlli=ko* *dud* *di-ya* *hɛ*
Nadya.F=ERG cat.F.SG=DAT milk.M.NOM give-PRF.M.SG be.PRES.3.SG
'Nadya has given milk to the cat.' Urdu

In both cases, the dative is associated with a more or less abstract goal. In (23) the 'cat' is the goal of the giving. In *experiencer* constructions as in (19b) or (24a–c), the experiencer of the event can be thought of as a kind of abstract goal (cf. Landau 2002).

(24) a. *nadya=ko* *skul* *ja-na* *paṛ-a*
Nadya.F.SG=DAT school.F.SG.OBL go-INF.M.SG fall-PRF.M.SG
'Nadya had to go to school.' Urdu
b. *nadya=ko* *ḍar* *lag-a*
Nadya.F.SG=DAT fear.M.SG.NOM be attached-PRF.M.SG
'Nadya was afraid.' Urdu
c. *nadya=ko* *kahani* *yad* *a-yi*
Nadya.F.SG=DAT story.F.SG.NOM memory come-PRF.F.SG
'Nadya remembered the story.' Urdu
d. *nadya=ne* *kahani* *yad* *k-i*
Nadya.F.SG=ERG story.F.SG.NOM memory do-PRF.F.SG
'Nadya remembered the story (actively).' Urdu

With psych predicates as in (24c) there is again an alternation with the ergative, though in this case it correlates with the use of the agentive 'do' in (24d) vs. the non-agentive 'come' in (24c). The dative use of *ko* is also governed by a combination of syntactic and semantic factors. It is restricted to indirect objects and subjects, but is subject to a coherent *goal* semantics in Urdu.

In DCT, there is thus explicit reference to a-structure concepts such as *goal* (datives) or *agent* (ergative) in addition to further semantic concepts such as specificity as part of the analysis of case. DCT recognizes case as an extremely complicated and complex part of the morphosyntactic and semantic interface, particularly as there are also some instances of case marking which appear to be tied to purely positional/structural or idiosyncratic factors. In DCT, therefore, examples such as the English adnominal genitive (e.g. *John's hat*) are analysed as purely positional/structural case.

DCT also assumes a notion of default case. Indeed, it can be observed that structural/positional case is often an instance of default case (the Elsewhere Case,

978–0–19–920647–6 05-Spencer-ch04 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 70 of 71 May 16, 2008 18:57

Kiparsky 1973). There are some languages which require that all NPs have case. For these languages default case satisfies something that has become known as the Case Filter (Rouveret and Vergnaud 1980). That is, if a given NP is not already associated with case due to some specification in some other part of the grammar, then default case assignment principles can apply. Typical cross-linguistic default cases are nominative or genitive. Default case only applies to the core grammatical relations subject and object. The other grammatical relations tend to involve some kind of specialized semantics and therefore do not involve defaults. Note that this is essentially the insight behind GB/Minimalism's postulation of exactly two structurally Case-marked arguments, which are associated with nominative and accusative case.

Finally, DCT also acknowledges that some instances of case marking are due to truly coincidental historical developments: no generalization can be made and the case marking has to be stated as an exception to the system. These instances of case marking are labelled *quirky*.

An example of truly idiosyncratic marking is shown in (25). Urdu requires that subjects of agentive transitive perfect verbs be marked with the ergative case. However, the verb 'bring' in (25) is a transitive verb with perfect morphology and the bringer can be unproblematically classified as an agent. But the subject 'Nadya' in (25) is nominative and not ergative.

(25) *nadya* *kıtab* *la-yi*
Nadya.F.SG.NOM book.F.SG.NOM bring-PRF.F.SG
'Nadya brought a book.' Urdu

The reasons for this aberrant case marking are not clear. Also, there are no other agentive transitive verbs which behave like this. Therefore this bizarre nominative case marking must be stipulated as part of the lexical entry of 'bring' (the nominative on objects is regular in Urdu).

On the whole, however, DCT assumes that such instances of quirky case are relatively rare. Rather, most case marking phenomena should fall under the rubrik of *semantic case*, that is, as being constrained simultaneously by both syntactic and semantic conditions. This stands in marked contrast to many theories of case, but does seem to reflect the rather unique cross-linguistic positioning of case at the interface between lexical and clausal semantics, syntax and morphophonology.

4.5 Incorporating Optimality Theory

The advent of Optimality Theory (OT, see Chapter 6) was observed with a great deal of interest in LFG and adopted into the theory very quickly (Bresnan 2000).

978–0–19–920647–6 05-Spencer-ch04 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 71 of 71 May 16, 2008 18:57

The input to an evaluation by OT constraints is assumed to be f-structure and c-structure pairings. The task of the OT constraints is to pick the most optimal pairing.

Several OT-LFG analyses exist with respect to case and they tend to fall in line with OT analyses of case in general. That is, they adopt the insights with respect to differential case marking advanced by Aissen (1999, 2003). Some examples are Sharma's (2001) analysis of Kashmiri case clitics, Deo and Sharma's (2005) analysis of typological and diachronic variation in Indo-Aryan ergative patterns, Asudeh's (2001) analysis of optional case patterns in Marathi. These analyses have all the advantages and disadvantages of the general OT analyses that are discussed in Chapter 6 of this volume. Asudeh's analysis, however, proposed an interesting extension to OT, namely, the availability of optionality in the evaluation of the most optimal candidate. In addition, Lee's (2001a,b; 2003) work on bidirectional OT and output–output correspondences provides a new perspective on Hindi and Korean case as well as word order freezing effects, which are due to identically case-marked arguments in a clause.

4.6 Summary

In sum, LFG contains a range of differing approaches to case marking. No account would claim to have developed a complete theory of case marking cross-linguistically. Rather, new ideas are continually being developed and tried out in light of new empirical evidence. This does not mean, however, that each account discards already established insights. Rather, each new account builds on the strong points of previous insights within LFG, but is simultaneously not willing to let itself be blinded by the previously established ideas.

CHAPTER 5

THE CASE TIER

A HIERARCHICAL APPROACH TO MORPHOLOGICAL CASE*

JOAN MALING

In this chapter I present a theory of case assignment based on a hierarchy of grammatical functions, as advocated in Zaenen, Maling, and Thráinsson (1985), Yip, Maling, and Jackendoff (1987), and Maling (1993), inter alia. I focus on the assignment of the syntactic cases nominative and accusative, singling out two phenomena: (i) nominative objects, and (ii) case alternations on adverbials of duration. A hierarchical as opposed to a structural analysis of nominative case provides a simple account for the occurrence of nominative objects with dyadic verbs which take oblique subjects, a phenomenon found in many genetically unrelated languages including Icelandic, Finnish, and Korean. Less common are languages in which certain adverbials are sometimes marked accusative, sometimes nominative, depending on other properties of the clauses in which they occur. I review a complex array of case patterns focusing on certain adverbials which are sometimes

*Acknowledgements: I am indebted to Urpo Nikanne, who happened to be in the right place (Brandeis) at the right time (the academic year 1989–90), for the generous help that made it possible for me to pursue the implications of the Case Tier analysis in Finnish. For Korean I am indebted to Soowon Kim and Jongsup Jun for years of patiently answering my questions about Korean morphosyntax. I thank the editors and an anonymous reviewer for useful comments on an earlier draft. Space limitations have occasionally made it necessary to ignore the subtleties of the more complex patterns.

marked accusative, sometimes nominative, depending on other properties of the clauses in which they occur. These complex patterns are difficult for any theory of case to handle because there is no one-to-one relation between morphological case and either syntactic position or grammatical function (see also Butt 2006, Spencer 2003). The data, drawn primarily from Finnish (Maling 1990, 1993), pose a serious empirical challenge to most theories, and to the best of my knowledge, these case alternations have not yet received any alternative formal account. The data support the view that nominative and accusative are not properties of the verb, but rather properties of the clause, as has been argued for by Kim and Maling (1998) and Sells (2004) on the basis of quite different phenomena.

5.1 Case-assignment principles

The fundamental question to be addressed is this: what is the relationship between morphological case and grammatical functions? In particular, what are the principles governing the assignment of the grammatical cases, NOM and ACC, and what is the domain to which the grammatical cases can be assigned? Both traditional grammars and recent generative studies on Case Theory recognize that case can be assigned in (at least) three different ways, which I will call semantic, lexical, and syntactic or grammatical case assignment. I distinguish grammatical case from true 'structural' case assigned to a particular syntactic position.

(1) a. GCASE: grammatical/syntactic case: NOM, ACC
b. LCASE: lexical case: e.g. DAT, GEN objects; oblique (i.e. non-nominative) subjects
c. SCASE: semantic/adverbial: accusative of time/duration, instrumental dative
d. Configurational/structural case: e.g. adnominal GEN (case assigned to [Spec, NP])

Grammatical or syntactic case is what is widely referred to as regular or 'default' case marking, which results in nominative subjects and accusative objects.[1] It is sensitive to surface grammatical relations, hence we find case alternations between NOM and ACC as a given nominal argument is passivized or raised (e.g. in the accusative plus infinitive construction, also known as Exceptional Case Marking).

Semantic case marking includes such traditional 'adverbial' categories as accusatives of time or duration, and instrumental datives. For the present purpose,

[1] The implications for ergative case marking systems will not be explored here (but see Yip et al. 1987 for tentative suggestions). In ergative languages, absolutive case seems to be 'default' in the sense of unmarked, but ergative is not referred to as a default case in the way that nominative is on subjects in nominative–accusative case systems.

this sort of invariant case marking on adverbials is uninteresting because it tells us nothing at all about the principles governing the assignment of grammatical case; for this, we must turn to languages where the case marking on adjuncts alternates.

Lexical case marking is an idiosyncratic property of a lexical item, assigned by a verb, preposition, or adjective. Lexical case is associated with a particular thematic role, and is assigned before arguments are associated with grammatical functions. This ordering accounts for the fact that passive and raising are 'case preserving' (Zaenen et al. 1985). Traditional grammar has noted the existence of fairly regular correlations between the case and thematic role of a given NP, but there are also many minimal pairs which show that morphological case cannot be predicted from the thematic role; lexical case marking is therefore assumed to be unpredictable and part of the lexical entry of a verb. Icelandic is notoriously rich in lexical case marking: hundreds of verbs govern dative in Icelandic (Maling 2002); Finnish and Korean exhibit significantly less lexical case marking, but I assume that the dative on the complements of the Korean verbs *kamsa-ha* 'thank', *kissu-ha* 'kiss', and *takao* 'approach' is lexical case.

Under the Case Tier hypothesis, the grammatical cases are assigned along a hierarchy of grammatical functions (GFs), where SUBJ is higher than OBJ. The primary motivation for a hierarchical account of case assignment comes from the existence of nominative objects, which are common cross-linguistically whenever the SUBJ is not marked nominative. A basic prediction of the Case Tier model is that the case marking on subject and object do not vary independently, but rather, lexical case marking on one argument has a concomitant effect on the syntactic case(s) assigned to the remaining verbal arguments, an effect I refer to as 'Case shift'. The Case-shift phenomenon is especially clear in Icelandic, where passives of *give*-type verbs also occur in the DAT NOM case frame: the retained object (the theme), which is accusative in the active sentence, is marked nominative in the passive voice:

(2) a.	*Jón* *gaf* *barninu*	DAT	[Lexical Case, pre-linked]
	Jon(NOM) gave the.child(DAT)	|	
	bókina	NP_1 $\not{NP}_2$	[GFs]
	the.book(ACC)		
b.	*Barninu* *var* *gefin*	NOM ACC	[Grammatical Case Tier]
	the.child(DAT) was given		
	bókin		
	the.book(NOM)		

Syntactic tests identify *barninu* as the grammatical subject of (2b) (see Zaenen et al. 1985, inter alia). On the other hand, passive verbs take unambiguously accusative complements in many languages including Classical Greek and Quechua. Ditransitive verbs with two accusative objects in the active voice take an accusative retained object, as illustrated in (3). English is another language of this type.

(3) a. *didáksomai* ***lógous*** Classical Greek
teach-1sg.pass.fut. speeches-ACC
'I shall be taught speeches' Euripides, Andromache 739
b. *can-ga* ***cuintu-ta*** (*wawa*) *villa-shca ca-rca-ngui* Imbabura Quechua
you-TOP story-ACC (child) tell-PASS be-PAST-2SG
'Were you told the story (by the child)?' (Jake 1985: 66, 17b)

5.2 What is the domain of the Case Filter?

What GFs can the grammatical cases map onto? A basic assumption of standard Case Theory in Government-Binding Theory (Chomsky 1981; Bobaljik and Wurmbrand, this volume) was that only verbal arguments fall within its domain, i.e. only verbal arguments are subject to the Case Filter. This is because Case Theory was assumed to explain NP-movement, and only verbal arguments undergo NP-movement (i.e. passive, raising). However, a wide variety of adverbials get assigned the same 'objective case' as verbal objects: measure phrases, cognate objects, and adverbials of duration and frequency. In Slavic languages, not only are such adverbials marked accusative, but just like objects, they get marked genitive under negation (see Babby 1980, Przepiórkowski 1999, inter alia).

5.2.1 Cognate objects

Many normally intransitive verbs can occur with a cognate object, an NP-complement whose head N is a result or state nominalization of the verb, as in *He died a horrible death* or *The lion roared a terrible roar*. For discussion of cognate objects within the generative literature, see Austin (1982) and Jones (1988). At issue is the status of cognate objects with respect to both the Case Filter and the Theta Criterion. Their status as grammatical objects is not clear, since they typically resist passivization; however, since Google turns up sentences like 'Smiles were smiled', there is clearly no absolute ban against passivizing them. Semantically they function like manner adverbials, and in some languages, they are case-marked as such. In Icelandic, for example, the cognate object to a strictly intransitive verb typically bears what might be called the 'instrumental dative':

(4) a. *Hann svaf* ***djúpum svefni.***
he slept deep sleep-DAT
'he slept a deep sleep'

978–0–19–920647–6 06-Spencer-ch05 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 76 of 87 May 16, 2008 18:57

b. *Hann dó* ***hræðilegum dauðdaga.***
he died horrible-DAT deathday-DAT
'he died a horrible death'

But in many unrelated languages, including Classical Arabic, Finnish, Latin, and Korean, cognate objects are marked accusative. What is the source of this accusative case? There are three possibilities to explore: (i) cognate objects bear an Inherent/Semantic Case appropriate to their adverbial interpretation; (ii) cognate objects bear syntactic case (typically ACC); or (iii) cognate objects lie outside the domain of the Case Filter, but in languages with overt morphological case, they get a default morphological case. Jones (1988) argues that the Case Filter applies only to argument NPs, and that cognate objects are adjunct-predicates rather than arguments. Cross-linguistic evidence shows that the syntactic case approach is more promising. Finnish, Russian, and Korean all have cognate nominatives:

(5) a. *Mne mečtalas'* ***jasnaja mečta.*** Russian
me-DAT dreamt clear dream-fem.SG.NOM
'I dreamt a clear dream'

b. *Chelsu-nun* ***yongkkwum-i*** *kkwu-ko siph-ta.* Korean
Chelsu-TOP dragon.dream-NOM dream-COMP want-DECL
'Chelsu wants to dream about a dragon' (lit. dream a dragon dream)

Cognate objects get marked nominative under the same conditions as true objects. DAT NOM is exactly the case frame one would expect for a verb with a dative experiencer in Russian, and NOM is what one would expect on the object in the *siph-ta* construction (Kim and Maling 1998). It is clear, then, that the term 'cognate accusative' is a misnomer: it is not a special function of the accusative to mark cognate objects. However they are licensed, cognate objects get whichever syntactic case would be expected given the case assigned to the other verbal argument(s). This indicates that even normally intransitive verbs can assign accusative case/objective case.

5.2.2 Adverbial adjuncts

How do adverbial adjuncts get case? In English, as in many languages, adverbials are typically PPs; even so-called adverbial NPs can plausibly be analysed as getting Case from a null P (Emonds 1987). In languages with overt morphological case, like Icelandic or German, adverbial NPs are said to bear accusative of time or duration, instrumental dative, etc. Such case-marking is invariant: unlike an accusative object, an accusative time adverbial in Icelandic or Latin does not exhibit case alternations depending on the transitivity or voice of the matrix verb. Hence this type of case marking is said to be 'semantic' or (within GB) 'inherent' case analogous to the lexical case marking on verbal arguments which is characterized

by case preservation under movement (Zaenen et al. 1985). This invariant case marking is uninteresting for the purposes at hand; more informative are languages like Finnish in which these adverbials exhibit case alternations. Certain classes of adverbials seem to get syntactic case in a surprising number of typologically diverse languages, including Finnish, Korean (Maling 1989), and Australian languages such as Warumungu (Simpson 1991).

There are two facts about this phenomenon to note. First, the same class of adverbials shows up again and again: adverbials of duration/frequency, measure phrases, cognate objects, suggesting the existence of a natural class.[2] Why should just this class of adverbials be 'object-like'? I won't try to answer this question – a task I happily leave to the semanticists – other than to note that co-occurrence with duration/frequency adverbials is one of the main tests for the four aspectual verb classes (Dowty 1991). Second, not only do these adverbials get the same ACC case assigned to verbal objects, but they exhibit the case alternations diagnostic of structural Case. In other words, they bear NOM case in exactly the same environments that verbal objects do.[3]

In the remainder of this chapter, I will explore the predictions of the Case Tier model for case assignment by investigating the assignment of grammatical cases to verbal objects and certain adverbial adjuncts, and the interaction of these cases with other sources of morphological case.

5.3 The case of Finnish objects

Finnish is said to have four grammatical cases: nominative, accusative,[4] genitive, and partitive. Accusative on objects alternates with partitive under negation. Partitive is considered the unmarked case on complements, while Accusative carries

[2] Note that grammatical case is sometimes found on other kinds of time adverbials, as in the following Warlpiri example cited by Spencer (2003), taken from Nordlinger (1998b, ex. (50)):

(1) Jalangu-rlu ka-lu-jana puluku turnu-ma-ni yapa-ngku
today-ERG PRES-3PL.S-3PL.O bullock(ABS) muster-CAUS man-ERG
'The people are mustering cattle today'

The adverbial *jalangurlu* 'today' is marked with ergative case because that is the case of the subject.

[3] Since adverbials do not undergo either passive or raising, this fact has obvious implications for case-theoretic accounts of NP-movement, which is supposedly triggered by the lack of Case.

[4] The nomenclature surrounding nominative objects in Finnish is a matter of impassioned controversy because of case syncretism in the nominal paradigm. An accusative singular NP bears either a zero-suffix (the so-called 'endingless accusative') or an *-n* suffix (historically *-m*), depending on the form of the matrix verb. The zero-suffix also occurs in a variety of systematically impersonal constructions, including passives and imperatives. Glosses in this paper reflect the claim that the accusative singular suffix is always *-n*, and the 'endingless accusative' is actually NOM.

with it a telic interpretation. This ACC/PART alternation distinguishes true objects and object-like measure phrases from object-like adverbials. ACC on objects alternates with NOM when the subject is (broadly speaking) unexpressed, as in imperatives, impersonals, and passives.

5.3.1 The case of adverbials

It is well-known that certain 'adverbials of amount' are treated like verbal objects for the purposes of case assignment in Finnish (Tuomikoski 1978). Measure phrases, cognate objects, and adverbials of duration and frequency get assigned the same case as objects, and like objects, get partitive case under negation. The significant fact is not that adverbials bear 'objective case', but that they exhibit the same case alternations between NOM and ACC that verbal objects do. This is illustrated in (6):

(6) a. (*Minä*) *viivyin matkalla* ***viikon***. Finnish
(I-NOM) stay-1SG trip-ADESS week-ACC
'I stayed on the trip a week'
b. *Siellä viivyttiin kokonainen* ***viikko***.
there-ADESS stayed-IMPF [whole week]-NOM
'We/they/one stayed there a whole week.'

If the matrix verb is finite and agrees with the (possibly null) NOM subject, then the adverbial is ACC; if the matrix verb is systematically impersonal, e.g. in the so-called passive, then the adverbial is NOM.[5] It is the case alternation which makes this look like grammatical case, and which distinguishes this from the use of accusative on certain types of adverbials in Icelandic, Russian, or Latin.

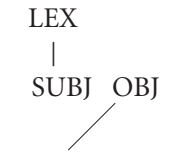
Let us assume that there is no appropriate semantic case in Finnish for duration/frequency adverbials. The adverbial NP will, therefore, have to get morphological case from some other source, namely, grammatical case from the syntactic Case Tier. Under the Case Tier hypothesis, NOM case marking on the object or adverbial is an automatic consequence of the Oblique case on the subject, which induces a shift of the syntactic case tier such that NOM appears on the object, if any, or on the adverbial of amount. The source of the oblique case marking on the subject in Finnish can be either a lexical property of the matrix verb as in (7), or the result of a

[5] A word of caution is in order. Certain adverbials e.g. *joka päivä* 'every day' are invariantly NOM for all speakers, whereas others like *pitkän aikaa* 'a long time' only vary between accusative and partitive, and do not alternate with NOM. Only adverbials which exhibit the NOM/ACC alternation are considered here. Similar variation is found in Korean: the adverbial *phyengsayng* 'entire life' is a fixed form in the sense that it takes ACC even when other duration adverbials take NOM (Maling, Jun, and Kim 2001, fn.2). For both Finnish and Korean, further research is needed to determine the extent to which these differences are semantically based, or simply idiosyncratic lexical properties of particular adverbial phrases. Finnish grammar books report that the language is still in a state of flux with respect to these issues (Lauri Karttunen, p.c.).

general rule assigning GEN configurationally to Spec position, as in (8b,c) with the modal *täytyy*.[6]

(7) a. *Lapsen oli* ***jano***
child-GEN be-pst thirst-NOM
'the child was thirsty'

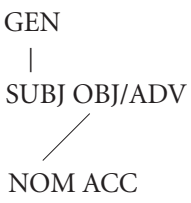
b. *Lapsella oli* ***kirja***.
child-ADESS be-PAST book-NOM
'a/the child had a book'



(8) a. *Lapsi luki* ***kirjan*** *illassa*
child-NOM read-3SG book-ACC evening-INESS
'the child read a/the/one book in one evening'

b. *Lapsen täytyy lukea* ***kirja***/**kirjan*.
child-GEN must read book-NOM/*ACC

c. *Lapsen täytyy lukea* ***koko päivä***/**päivän*
child-GEN must read all day-NOM/*ACC



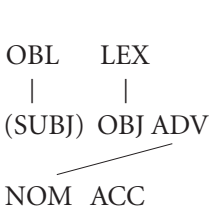
Following Vainikka (1989), I assume that GEN is a purely structural or 'configurational' case in Finnish, assigned to specifier position for all lexical categories.

5.3.2 Combining objects and adverbials

We have seen that NOM from the Syntactic Case Tier can shift over to the OBJ, if any, or if there is no OBJ, onto certain adverbials. The obvious question is what happens if both DO and ADV co-occur? Lexical case marking on an object has exactly the same effect as lexical case marking on a subject: syntactic Case shifts over to Adverbial, exactly as expected under the Case Tier Hypothesis. This is illustrated in (9) for the verb *luotta* 'to trust', which governs Illative (ILL) case on its object:

(9) a. *Kansa luotti Kekkoseen* ***vuoden***.
people-NOM trust-pst-3SG Kekkonen-ILL year-ACC
'people trusted Kekkonen for a year'

b. *Kekkoseen luotettiin* ***vuosi***.
Kekkonen-ILL trust-IMPF year-NOM
'Kekkonen was trusted for a year'



[6] See Maling (1993) for arguments against claiming that the modal *täytyy* assigns GEN lexically rather than structurally.

c. *Kansan täytyy luottaa Kekkoseen **vuosi**.*
people-GEN must trust Kekkonen-ILL year-NOM
'the people must trust Kekkonen for a year'

It is striking that if both SUBJ and OBJ bear an oblique case, NOM from the syntactic Case Tier shifts over not just one, but two NPs, all the way to the Adverbial NP, as illustrated in (9c).

What happens if we try to combine a syntactically case-marked object and a syntactically case-marked adverbial? Under the Case Tier Hypothesis, we expect to be able to generate the case pattern OBL NOM ACC whenever the subject bears an oblique case. The case pattern GEN NOM ACC is ruled out for agentive verbs where the use of grammatical case marks telicity and the adverbial is durational. But under the Case Tier hypothesis, this case pattern is predicted to occur, and indeed it does. Choosing a non-agentive predicate with oblique subject as in (10a,b) or a frequency adverbial as in (10c) makes the case pattern acceptable.

(10) a. *Lapsen oli jano*
child-GEN be-PAST thirst(NOM)
***yhden illan**/*yksi ilta*
[one evening]-ACC/*NOM
'the child was thirsty for one evening'

LEX
|
SUBJ OBJ ADV
/ /
NOM ACC

b. *Lapsella oli kirja **koko päivän**/*päivä.*
child-ADESS be-PAST book-NOM [all day]-ACC/*NOM
'a/the child had the book all day'

c. *Lapsen täytyy lukea kirja **kolmannen kerran**/***kolmas kerta**.*
child-GEN must read book-NOM [third time]-ACC/*NOM
'the child must read the book for a third time'

Another way of exemplifying the expected case pattern is to find a transitive verb where the use of structural case does not mark completion and hence the meaning is compatible with a duration time adverbial (see Maling 1993).

5.3.3 Case spreading

Finnish has no double accusative verbs. Ditransitive verbs have only one ACC object, while the other complement bears a locative case (allative, ablative, adessive, or elative), which is usually semantically predictable but may be idiosyncratic (Vainikka 1989: 326). One possible hypothesis would be to attribute the lack of double accusative verbs to the standard case-theoretic assumption that any case assigner has at most one syntactic case to assign[7]; Li (1990) observes that duration/frequency

[7] This is not universally true, however. Double accusative ditransitive verbs can be found in several unrelated languages including Quechua and Classical Greek. Icelandic has no double ACC ditransitive verbs, but does have a number of ditransitive verbs which assign lexical case to both objects, e.g. *lofa*

adverbials cannot co-occur with a postverbal object in Chinese, and suggests that both NPs compete for a single syntactic accusative case. In Finnish, however, unlike Chinese, an accusative object and an accusative adverbial can co-occur, as illustrated in (11).

(11) *Minä luin kirjan* ***kolmannen kerran.*** SUBJ OBJ ADV
I-NOM read book-ACC [third time]-ACC | |/
'I read the book for a third time' NOM ACC

This indicates that Finnish has Case-spreading. Under the Case Tier hypothesis, we expect NOM to be assigned to the first available NP on the GF hierarchy, and ACC from the Case Tier to spread rightwards, associating with any available NP(s). This is exactly what happens when there is more than one adverbial NP in a clause: only the first adverbial gets NOM in impersonal constructions, and the others get ACC.

(12) a. *Käveltiin* ***koko talvi****/*talven*
walked-IMPF [whole winter]-NOM/*ACC
'there was walked the whole winter'
b. *Käveltiin kilometri* ***koko talven****/*talvi*
walked-IMPF kilometer-NOM [whole winter]-ACC/*NOM
'there was walked a kilometer the whole winter'

It seems unlikely that the first adverbial gets NOM because it has actually been 'promoted' to OBJ, given that the same NOM case marking occurs in the presence of a lexically case-marked object, as in (13).

(13) *Kekkoseen luotettiin* ***kolmas kerta.***
Kekkonen-ILL trust-IMPF [third time]-NOM
'Kekkonen was trusted for a third time.'

What exactly is meant by 'the first' adverbial? The linear order of the adverbial phrases is, not surprisingly, irrelevant. As illustrated in (14b–d), it is the duration phrase which gets marked NOM (*yksi vuosi*)[8] while the freqency phrase gets ACC (*yhden kerran*).

(14) a. *Kekkoseen luotettiin* ***yksi kerta***
Kekkonen-ILL trust-IMPF [one time]-NOM
'Kekkonen was trusted once'
b. *Kekkoseen luotettiin* ***yksi vuosi*** DUR-NOM>FREQ-ACC
Kekkonen-ILL trust-IMP [one year]-NOM
yhden kerran
[one time]-ACC
'Kekkonen was trusted for one year once'

'to promise' which governs two dative objects. For a discussion of the different case frames for Icelandic verbs, see Yip et al. (1987) and Maling (2002).

[8] Urho Kekkonen is a former president of Finland who was elected to several six-year terms of office. Since cardinal numbers other than *yksi* 'one' are invariant, it is impossible to tell the case of *kuusi* 'six', So I have changed the examples to the pragmatically less natural 'one year.'

c. *Kekkoseen* *luotettiin* *yhden kerran* FREQ-ACC>DUR-NOM
Kekkonen-ILL trust-IMPF [one time]-ACC
yksi vuosi
[one year]-NOM

d. **Kekkoseen* *luotettiin* *yksi kerta* *FREQ-NOM>DUR-ACC
Kekkonen-ILL trust-IMPF [one time]-NOM
yhden vuoden
[one year]-ACC

The fact that linear order does not affect the case-marking of the adverbials suggests that the Case Tier is defined over GFs and that the GF hierarchy must distinguish between the different classes of adverbials, to ensure that a duration adverbial takes precedence over the frequency adverbial in getting NOM case.

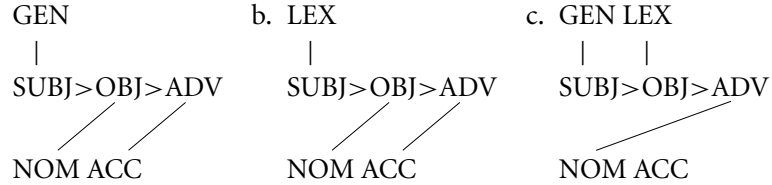
5.3.4 Summary of the Case Tier

The fundamental assumption of the Case Tier Hypothesis is that NOM and ACC form a syntactic case tier which is mapped onto the NPs of a sentence one-to-one, left-to-right, by principles analogous to those of autosegmental phonology. The basic predictions of this Case-Tier Hypothesis are summarized in (15):

(15) a. NOM is assigned before ACC.
b. Only one XP can get assigned NOM; any remaining NPs get ACC via case-spreading
c. Which XP gets NOM reflects the hierarchy of GFs, where SUBJ>OBJ>DUR>FREQ

On the further assumption that NPs already bearing morphological case, however assigned, are unavailable for mapping onto the syntactic Case Tier, the presence of such case marking will induce the by now familiar Case Tier shift. Because Finnish has lexical case marking and includes adverbials in the domain of the syntactic Case Tier, Finnish provides convincing evidence of Case Tier shift from various sources of case marking. These are sketched in the diagrams in (16), where ADV stands for any of the various adverbial functions lower than OBJ.

(16) Case Tier shifts



If there is no extra source of case marking, then the ACC spreads to any additional NPs. As we have seen, all of these case patterns are attested in Finnish.

978–0–19–920647–6 06-Spencer-ch05 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 83 of 87 May 16, 2008 18:57

5.4 Non-finite complements

Thus far, we have considered only monotransitive verbs. Let us now consider VP complements which occur as arguments of a matrix verb. Finnish has both clausal (IP) and nonclausal (VP or V′) complements. Since finite complements may be assumed to have their own Case Tier, only the non-finite complements are of interest here. Of particular interest is the fact that the alternation between NOM and ACC on objects in a single clause extends to the objects of infinitival complements. The choice between NOM or ACC case depends on the matrix verb: if the matrix V agrees with a NOM subject, the infinitival object is ACC, as in (17a); if the matrix verb is such that its object would occur in NOM (i.e. passive, imperative, and in certain impersonal verbs), then the infinitival object must be marked NOM, as in (17b) (Vainikka 1989: 255, ex. 8a,b).

(17) a. *Pekka muisti käydä ostamassa* ***solmion****/*solmio*
Pekka-NOM remembered go-TA buy-INF tie-ACC/*NOM
'Pekka remembered to go buy a tie'
b. *Muista käydä ostamassa* ***solmio****/*solmion!*
remember go-TA buy-INF tie-NOM/*ACC
'Remember to go buy a tie!'

The basic problem, then, is to account for why the embedded VP 'inherits' the case-assigning properties of the matrix V. It is clear that these non-finite complements cannot be analysed as having PRO subjects which must agree in Case with their matrix controller. This would not give the desired result within the Case Tier framework, since whatever decision we made about assigning Case to PRO, it would not vary depending on the matrix verb, and we would not get the desired alternations in Case. Even if the matrix V is [+completed], its agreement features determine the grammatical case (NOM vs. ACC) on the object of the infinitival complement:

(18) a. *Maija pyysi Jukan lukemaan* ***kirjan.***
Maija-NOM asked Jukka-ACC read-MA-ILL book-acc
'Maija asked Jukka to read the book' (Vainikka 1989: 267, ex. 18b)
b. *Pyydä Jukka lukemaan* ***kirja****/*kirjan!*
ask Jukka-NOM read-MA-ILL book-NOM/*ACC
'Ask Jukka to read the book!' (Vainikka 1989: 267, ex. 19)

If the infinitival object is marked NOM, ACC can be realized on an adverbial NP:

(19) *Pyydä Jukka lukemaan kirja* ***kolmannen kerran.***
ask Jukka-NOM read-MA-ILL book-NOM [third time]-ACC
'Ask Jukka to read the book for a third time!'

If such adverbials get case from the syntactic Case Tier, as I have argued here, we see that ACC is still available even in those constructions which form the domain of

'nominative objects' in Finnish. This is as expected under the Case Tier Hypothesis: the availability of NOM and ACC is entirely independent of the assignment of GEN to [Spec,VP].

Adapting Vainikka's analysis of infinitival complements to the Case Tier Hypothesis, let us assume that such VP-complements are transparent to case assignment from the matrix clause. In other words, unlike IPs, VPs are not a barrier to case assignment. Since these infinitival complements function as arguments of the matrix verb, let us assume that they are assigned syntactic case in the same way that NP-complements would be. However, since a VP is not itself capable of bearing Case, the Case feature is transmitted from the embedded VP to its head (as suggested, for example, by Rouveret and Vergnaud (1980) for Romance causatives).[9] This gives the appearance that structural case can percolate down into the infinitival VP-complement to any 'receptive' NP.

Although Case Tier shift easily accounts for the ACC case marking on an adverbial following a NOM object, examples like (19) pose a different problem for the Case Tier Hypothesis, since NOM is assigned to two different NPs, both the matrix object and the infinitival object. How can this spreading of NOM be reconciled with the apparent Case-shift of ACC onto the adverbial NP? I suggest that the solution to this problem lies in the following descriptive generalization.[10]

(20) All **internal** arguments of a predicate must get the same grammatical case.

The clearest evidence for this generalization comes from Korean, which is well-known for having a variety of 'multiple Accusative' constructions. When such sentences are passivized, all of the accusatives become nominative, as illustrated above:

(21) a. *Cheli-ka* ***Mary-lul*** ***panci-lul*** ***senmul-ul*** *ha-ess-ta* ACTIVE
Cheli-NOM Mary-ACC ring-ACC gift-ACC do-PAST-DECL
'Cheli presented Mary with a ring'
b. *Mary-ka* ***panci-ka*** ***senmul-i*** ***toy-ess-ta*** PASSIVE
Mary-NOM ring-NOM gift-NOM become-PAST-DECL
'Mary was presented with a ring'

Now consider different case-marking properties of the so-called lexical passive, illustrated below in conjunction with the Inalienable Possession Construction. Unlike the syntactic passive, the lexical passive is ambiguous between a case-absorbing

[9] The suffixes marking the infinitival complements are glossed here simply as INF, obscuring the fact that the VP-complement is itself assigned a locative case which is realized on the infinitival head, illative in the case of the Finnish verb for 'ask'. This is the same case that would be assigned to a corresponding PP-complement (Nikanne 1990). This means that syntactic case percolates down through a complement which is itself assigned lexical case.

[10] If this suggestion is on the right track, it indicates that the Case Tier maps onto both direct and indirect objects at once; in other words, the GF hierarchy is not articulated into two object functions, at least for the purposes of assigning syntactic case. Different internal arguments can, of course, bear different morphological cases if other sources of case marking are available in a language, e.g. lexical case marking.

passive (22a) and a verb which is semantically passive but nonetheless an accusative case assigner (22b):

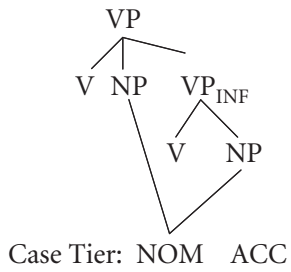
(22) a. *Yumi-ka* ***phal-i*** ***oynccok-i/*ul*** ***kkuth-i/*ul***
Yumi-NOM arm-NOM left-side-NOM/*ACC end-NOM/*ACC
cap-hi-ess-ta.
catch-Pass-PAST-DECL

b. *Yumi-ka* ***phal-ul*** ***oynccok-ul/*i*** ***kkuth-ul/*i***
Yumi-NOM arm-ACC left-side-ACC/*NOM end-ACC/*NOM
cap-hi-ess-ta.
catch-pass-PAST-DECL
'Yumi was caught by the arm on the left side at the end'

Superficially the object(s) of a lexical passive can be either NOM or ACC; however, if there is more than one object, their case marking cannot vary independently. Since only one NP can be the surface subject, the remaining NPs must all be internal arguments, and the internal arguments necessarily share the same syntactic case. See Maling and Kim (1992) for further discussion.

Returning to Finnish, then, recall our assumption that infinitival VP-complements are transparent to case assignment from the matrix clause, and get assigned syntactic case in the same way that an NP-complement would. For simple clauses, there is no overt evidence for the generalization that all internal arguments bear the same syntactic case; as noted above, any verb that subcategorizes for two NPs has only one ACC object, while the other complement bears a locative case (ALL, ABL, ADESS, ELA) which is usually semantically predictable, but may be idiosyncratic (Vainikka 1989: 326). However, when the locative complement is infinitival, we do see overt evidence that the same syntactic case is assigned to both complements, since the same case is realized on both matrix and infinitival object, as sketched in (23).[11]

(23)



What we want to say is that verbs which take infinitival complements are not exceptional, but have exactly the same case-assigning properties as other verbs of the language. This phenomenon is reminiscent of causativization: cross-linguistically, the output behaves like a single verb/complex predicate with respect to case assignment.

[11] See Vainikka (1989: 260) for discussion of the phrase structure of (object-)control verbs.

Two alternative analyses come to mind: merger (complex predicate formation), or alternatively, we might assume that the Case Tier is associated only with IP, so that when the infinitival complement is a bare VP, only one Case Tier is available. Whichever analysis is correct for Finnish, as internal arguments of the matrix verb, both NP and the infinitive itself must be assigned the same syntactic case in accordance with the generalization in (20).

When NOM is assigned to the internal arguments, the unassigned Accusative is still available, and can shift to an Adverbial. Adverbials can modify either the matrix or the embedded infinitive:

(24) *Pyydä Jukka* ***kolmannen kerran*** *lukemaan kirja.*
Ask Jukka-NOM third-ACC time-ACC read-INF book-NOM
'Ask Jukka for the third time to read the book!' or
'Ask Jukka to read the book for the third time!'

Further evidence for this analysis comes from verbs that assign lexical case to their objects. In (25), the embedded infinitive *vastaa* 'to answer' takes an ILL object. If the VP-complement supplied its own Case Tier, we would expect NOM from the Case Tier to shift over to any Adverbial; however, this is not what happens.

(25) *Oppilas pakotettiin* ***kolmannen kerran/****kolmas kerta vastaamaan*
pupil-NOM force-IMPF third time-ACC/*NOM answer-INF
kysymykseen ***tunnin/****tunti.*
question-PL-ILL hour-ACC/*NOM
'The pupil was forced for the third time to answer questions for an hour'

The case on the adverbial modifying the infinitive is determined not by the case-assigning properties of the infinitive, but by the properties of the matrix V. Since the matrix object is NOM, an adverbial modifying the embedded infinitive must be ACC. However, if the matrix object is not NOM but PART (implying resistance to the pressure), then NOM remains available for adverbials, as illustrated in (26):[12]

(26) *Oppilasta pakotettiin kolmas kerta vastaamaan kysymykseen*
pupil-PART force-IMPF [third time]-NOM answer-INF question-PL-ILL
tunti.
hour-NOM
'they tried to force the student for the third time to answer questions for an hour'

[12] When the matrix object is marked PART, the adverbials can also be marked ACC, as illustrated in (ii):

(1) Oppilasta pakotettiin kolmannen kerran vastaamaan kysymykseen tunnin.
pupil-PART force-IMPF [third time]-ACC answer-INF question-PL-ILL hour-ACC
'they tried to force the student for the third time to answer questions for an hour'

This unexpected fact may indicate that PART is being analysed not as a lexical case, which would induce the Case Tier shift seen in (26), but as an 'overlay' case (Yip et al. 1987) masking a covertly assigned NOM.

978–0–19–920647–6 06-Spencer-ch05 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 87 of 87 May 16, 2008 18:57

What is especially telling is the fact that a matrix adverbial and an embedded adverbial tend to agree in case: they are either both NOM or both ACC. This is as expected if the entire sentence is treated as a single domain with respect to the Case Tier. Of course, if the matrix verb is finite, then NOM on the adverbial is no longer possible.

In this section, I have argued that clauses with VP-complements are treated as one domain for the Case Tier: matrix and infinitival objects must share the same syntactic case, being either both nominative or both accusative.

5.5 Conclusion

The complex array of facts presented in this chapter receives a simple and elegant account under the hypothesis that the so-called 'structural Cases' NOM and ACC are assigned neither to particular configurational positions nor to particular grammatical functions, but along a hierarchy of grammatical functions. It is hoped that future research will lead to an account of the intriguing cross-linguistic variation[13] while maintaining the basic insight into the source of the observed case alternations provided under the Case Tier Hypothesis.

[13] Unlike Finnish, Korean allows multiple nominatives in a single clause, including D/F adverbials. This suggests that languages can choose whether to distinguish between adjuncts and internal arguments in the Case Tier hierarchy: Finnish does, Korean does not. Maling, Jun, and Kim (2001) show that it is also necessary to refer to the semantic distinction between states and events. Many details remain to be worked out, not least the way the case-marking on frequency adverbials in Korean fits into the picture (see e.g. Kim and Maling 1993). Most problematic is the fact that a NOM frequency adverbial can co-occur with an ACC duration adverbial, the very pattern disallowed in Finnish.

CHAPTER 6

CASE IN OPTIMALITY THEORY

HELEN DE HOOP

In Optimality Theory (Prince and Smolensky 2004), a grammar consists of a set of constraints which are violable and typically conflicting. That these constraints are violable implies that an output of an input–output mapping can never be rejected because it violates certain constraints or too many constraints. An output can only be rejected if there is a better output available (which means an output that does a better job in satisfying the complex set of constraints). Languages share an important subset of these constraints, which is why we may call these constraints 'universal'. However, individual languages rank these universal constraints differently in their language-specific hierarchies. Thus, one constraint can have very strong (categorical) effects in one language that ranks the constraint high in the hierarchy, while in another language the same constraint has hardly any effects, due to its low ranking. When a set of constraints is identified, the possible rankings of these constraints will generate the possible types of languages that the theory predicts. Thus, Optimality Theory (OT) provides a general means for constructing particular grammars from weighting certain constraints, and for generating theoretical typologies of possible languages from the same constraints.

6.1 From argument prominence to case assignment

The very first application of Optimality Theory in syntax happened to be on the subject of case (Legendre et al. 1993). Legendre et al. (1993) use a small number of constraints to formally account for the case patterns in different types of sentences in nominative–accusative, ergative–absolutive, and active/stative languages. The constraints apply to an input that is labelled for type of argument (agent, patient) as well as for abstract 'prominence' of the arguments (high-prominence, low-prominence). In the output nominative and ergative case are labeled C_1, absolutive and accusative case are labelled C_2, and oblique case as well as implicit arguments are labelled C_4. Probably the most important result of the approach by Legendre et al. (1993) is the fact that prominence demotions in the input drive voice alternations. Passives arise from an input consisting of a low-prominence agent and a high-prominence patient, written as *aP*. The constraint that 'low-prominence arguments do not receive one of the core cases C_1 or C_2' accounts for the fact that C_4 is assigned to the agent argument in a passive construction. That is, the demoted agent either gets an oblique case or is omitted altogether. This is illustrated in the simplified tableaux 6.1 and 6.2 below, with only three constraints (high-prominence arguments receive C_1; low-prominence arguments do not receive C_1 or C_2; some argument receives C_2), and their respective rankings in nominative–accusative and ergative–absolutive languages (for the full set of constraints, see Legendre et al. 1993).

Tableau 6.1. Derivation of a passive in a nominative–accusative language

Input: aP	X→ C_1	x → ¬ C_{12}	α → C_2
a_1P_2	*!	*	
a_2P_1		*!	
☞ a_4P_1			*
a_4P_2	*!		

When we have *Ap* as an input, we get the antipassive constructions A_1p_4 and A_2p_4 as optimal outputs in the two types of languages. As illustrated in the tableaux above, in a nominative–accusative language, the constraint which requires

Tableau 6.2. Derivation of a passive in an ergative–absolutive language

Input: aP	x → ¬ C_{12}	α → C_2	X → C_1
a_1P_2	*!		*
a_2P_1	*!		
a_4P_1		*!	
☞ a_4P_2			*

'high-prominence arguments to receive C_1' outranks the constraint that states that 'some argument receives C_2', whereas this is the other way around in ergative–absolutive languages. The ranking of these two constraints does not influence the outcome of transitive inputs (*AP*) which get A_1P_2 as the optimal output in all types of languages (in fact, there are more constraints involved, linking agents to C_1 and patients to C_2). But the ranking does influence the optimal output of passives, antipassives, and intransitive inputs, i.e. one-place predicates. The reader can verify that in nominative–accusative languages, the subject of an intransitive gets C_1 (nominative), while in ergative–absolutive languages, where the ranking is the other way around, the subject of an intransitive gets C_2 (absolutive).

In languages that have morphological case, subject and object cases are often structurally assigned, i.e. partly independent of the argument roles (agent, patient) and the high or low prominence of the arguments. The approach of Legendre et al., however, suggests that some independent prominence distinction will in fact trigger the occurrence of certain types of case, and in relation with that, the occurrence of grammatical voice. Sells (2001) points out that in practice, determining the input prominence of arguments is extremely difficult, one reason being that the overt expression of prominence distinctions not only involves case and voice marking, but also word order and intonation. That is, languages clearly differ in what the inventory of their given case and voice systems is and they may actually fail to represent given prominence distinctions in the input faithfully in case and voice morphology. Also, Legendre et al. assume that the agent and patient argument in an active (transitive) sentence are equally prominent in the input, and only differ in prominence in passive and antipassive constructions. Sells (2001), however, argues that in a canonical transitive sentence the agent (or subject) is already more prominent than the patient (or object), although it is 'much more prominent' in an antipassive construction. In fact, Aissen (1999, 2003) takes the difference in prominence between the subject and the object of a canonical transitive clause to be the driving force behind differential case marking patterns, as we will see below.

Obviously, differential case marking patterns constitute a problem for the account of Legendre et al. It turns out that the constraints on case marking are often sensitive to additional semantic restrictions, often related to the prominence of the arguments, that specify the conditions under which one or the other case marker is associated with the argument (de Hoop 1999; Stiebels 2000; Woolford 2001). This can be witnessed for transitive sentences, whose input should always be *AP* according to Legendre et al., but the problem is even obvious with intransitive inputs.

Another problem is that if the two arguments of a transitive clause are equal in prominence, we would expect the optimal output of a transitive input in a nominative–accusative language to be $A_1 P_1$, since the constraint 'high-prominence arguments receive C_1' is ranked high in these languages. An extra constraint is introduced by Legendre et al. to exclude the occurrence of the two arguments bearing the same structural case in the output, thus $^*A_1 P_1$, $^*A_2 P_2$. Although Legendre et al. (1993) do not treat any violations of this additional constraint (they introduce it as an absolute constraint or some kind of an extra assumption), it actually seems to work as a violable constraint as well, as recognized in subsequent work in OT syntax by e.g. Wunderlich and Lakämper (2001) and Anttila and Fong (2000), briefly discussed in the next section.

6.2 Uniqueness

Certain ungrammaticalities in language have been characterized in terms of a ban on identical morphological elements close to each other at some level of representation.

Anttila and Fong (2000) argue that the interaction between two constraints, one syntactic in nature, the other one semantic in nature, determines the choice between partitive and elative case in Finnish partitive constructions. The semantic constraint requires elative case to occur on quantitatively determinate noun phrases. Elative case is the marked case in a Jakobsonian markedness opposition, partitive case the unmarked. This is in accordance with the status of the partitive as the unmarked case for direct objects, argued for by Vainikka (1989) and Vainikka and Maling (1996).

The syntactic constraint that interferes with the semantic one is an instantiation of the constraint introduced above, called Uniqueness by Wunderlich and Lakämper (2001) as it prohibits the assignment of the same case twice. In the Finnish example, Uniqueness prohibits assignment of the same case to the head and the complement of a nominal constituent. In the constructions Anttila and Fong discuss, elative case is blocked within a noun phrase if the entire construction

978–0–19–920647–6 07-Spencer-ch06 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 92 of 101 May 16, 2008 18:57

bears elative case. The same applies to partitive. This is illustrated in (1) and (2). The verb *tulla* 'become' assigns elative case to its subject, hence the embedded noun must occur in the partitive:

(1) **Kolmasosa-sta miehi-stä tuli munkkeja.*
one-third-ELA men-ELA became monks

(2) *Kolmasosa-sta miehi-ä tuli munkkeja.*
one-third-ELA men-PART became monks
'One third of the men became monks.'

The verb *rakastaa* 'love' on the other hand takes a partitive object. Therefore, the embedded noun occurs in the elative:

(3) **Anders rakastaa tätä osa-a Helsinki-ä.*
Anders loves this part-PART Helsinki-PART

(4) *Anders rakastaa tätä osa-a Helsingi-stä.*
Anders loves this part-PART Helsinki-ELA
'Anders loves this part of Helsinki.'

So, in Finnish, UNIQUENESS which prohibits twice the same case on the head and the complement within a noun phrase appears to be stronger than the syntactic-semantic constraint that elative case correlates with quantitative determinacy.

Another manifestation of UNIQUENESS is found in languages that prohibit two arguments of a predicate to be identically case marked (cf. Wunderlich and Lakämper 2001, Maling 2001). Wunderlich and Lakämper (2001) extensively discuss these two different types of strategies in language as well as complicating interactions with yet other conditions in different varieties of Quechua. The basic idea is illustrated by the following two simple examples, example (5) from Huanca, where UNIQUENESS is violated, and example (6) from Cochabamba (cf. van de Kerke 1996), where UNIQUENESS is satisfied:

(5) *wamla-kaq-ta yaku-kta qu-y.*
woman-the-ACC water-ACC give
'Give water to the woman!'

(6) *ñuqa-man t'anta-ta qu-wa-nqa.*
1SG-DIR bread-ACC give
'S/he will give me bread.'

Wunderlich and Lakämper (2001) argue that in (5) UNIQUENESS is violated in favour of a constraint that requires objects to be marked with accusative case. Obviously, whereas the latter constraint wins in (5), UNIQUENESS wins in (6). We may hypothesize that both constraints exist in both dialects of Quechua. The difference is the mutual ranking between the constraints in Huanca and Cochabamba. Thus, the examples in this section show that conflicts between constraints may

978–0–19–920647–6 07-Spencer-ch06 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 93 of 101 May 16, 2008 18:57

lead to various kinds of outcomes, dependent on the relative strength (or order) of the constraints involved. Languages vary with respect to their weighting of the constraints.

6.3 Faithfulness

In Wunderlich (1997), Stiebels (2000), and Kiparsky (2001), an abstract level of representation is used for the characterization of argument properties. Argument roles are decomposed into primitive features such as for example *[+hr, −lr] = [there is a higher but not a lower role in the argument structure]*. Case assignment is determined by optimizing the correspondence between this abstract level of representation and the morphosyntactic output, subject to a set of ranked constraints. This means that in order to get the right case output, certain assumptions about the input sometimes must be postulated. Consider for example the object case alternation in Finnish and its concomitant meaning effect:

(7) *Anne joi maido-n.*
Anne drank milk-ACC
'Anne drank (up) the milk.'

(8) *Anne joi maito-a.*
Anne drank milk-PART
'Anne drank (some) milk.'

Kiparsky (2001) claims that at the (abstract) input level, the argument *milk* in (7) is neither the highest argument nor the lowest argument of the argument structure. The argument *milk* in (8), however, gets a different characterization. It is not the highest argument of the argument structure, but it is the lowest argument. At first sight, the object is the lowest argument in both (7) and (8), the subject being the highest one, and no other argument visible. In order to motivate the postulated difference in argument structure between the accusative and the partitive object in Finnish, Kiparsky claims that there is an extra 'result role' present in the argument structure of (7) that is absent in the one of (8). This result role is supposed to trigger the difference in object case marking between (7) and (8).

This is reminiscent of approaches such as Bittner and Hale's (1996b) account of ergative subjects of a certain class of intransitives in Hindi. An example of this type of differential subject marking in Hindi is given in (9) and (10) below.

(9) *Raam-ne jorse cillaayaa*
Raam-ERG loudly shouted
'Raam shouted loudly (volitionally)'

(10) *Raam jorse cillaayaa*
Raam loudly shouted
'Raam screamed loudly'

Bittner and Hale (1996b) analyse these intransitives as underlyingly transitive, comparable with cognate object constructions, such as *Raam screamed a loud scream*. That is, they also postulate an extra argument role at the input level, but the correlation with volitionality in the interpretation receives no direct explanation.

A similar alternation is found in Urdu which shows a volitionality difference between ergative and dative subjects in transitive sentences such as (11) and (12) below (Butt and King 2004). Wunderlich and Lakämper (2001) also account for this difference in volitionality in terms of a difference in the input argument structure.

(11) *anjum-ne xat lik^{h}naa hai.*
Anjum-ERG letter write is
'Anjum wants to write a letter.'

(12) *anjum-ko xat lik^{h}naa hai.*
Anjum-DAT letter write is
'Anjum has to write a letter.'

Wunderlich and Lakämper (2001) assume that the copula in these sentences assigns the feature 'there is a lower role' to the highest argument if it means 'want', while it assigns the feature 'there is a higher role' if it means 'be obliged'. Intuitively then, the higher role in (12) is the implicit argument that is in control of the action. That is, someone or something forces Anjum to write a letter.

So, in this type of approach, a difference in abstract argument structure is postulated in the input, in order to account for the difference in morphological case outputs. Thus, the different features that appear to affect case assignment cross-linguistically are derived from features related to the relative position of the argument in the argument structure. However, the input–output representations are very closely related, which means that the argumentation for having a certain argument structure in the input is partly *driven* by the data that it should *explain*.

Another problem for this type of analysis is the existence of different types of differential case marking, such as the partly independent case alternations for subjects and objects (i.e. independent of the case realized on the other argument) we observe in Hindi/Urdu, where the case on the subject may change independently of the case alternation on the object. An example is given in (13)–(14) (de Hoop and Narasimhan 2005):

(13) *wo ek laD.kaa / ek laD.ke-ko dekhtaa hae.*
he.NOM one boy.NOM / one boy-ACC sees
'He sees a boy / the boy.'

978–0–19–920647–6 07-Spencer-ch06 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 95 of 101 May 16, 2008 18:57

(14) *us-ne ek laD.kaa / ek laD.ke-ko dekhaa.*
he-ERG one boy.NOM / one boy-ACC saw
'He saw a boy / the boy.'

The *subject* case alternation in (13) and (14) is dependent on the perfectivity of the transitive sentence. Ergative case is assigned in the context of a perfective verb phrase. The *object* case alternation correlates with animacy and/or specificity of the object: accusative case marks a human or specific object. As one can witness, all four logical case combinations are possible on the subjects and objects of the sentences above, the case alternations do not depend on each other. There are many languages that show case alternations, such as differential object marking (DOM) or differential subject marking (DSM) or both. These alternations are often, yet not always, related to a difference in prominence of the arguments. Therefore, the approach of Legendre et al. (1993) in which prominence distinctions play a role in the input to case assignment might offer a solid basis for extension such that it can deal not only with voice but also with independent case alternations that are triggered by prominence distinctions. One highly influential OT approach that is developed to deal with argument prominence and differential case marking is Aissen's harmonic alignment account of DSM and DOM (1999, 2003).

6.4 Markedness

As Aissen (1999, 2003) argues, sometimes the function of case marking is not to mark all subjects or all objects of a transitive clause, but rather to mark those that are 'less prototypical' subjects or objects. In other words, only those objects that have certain characteristics of canonical subjects (such as being animate or specific) are case-marked, or only those subjects that have characteristics of objects (such as being a noun phrase rather than a pronoun or being a third person pronoun rather than a first or second one) (Silverstein 1976, Comrie 1989). The functional motivation behind this type of differential case marking would be to avoid ambiguity as to what is the subject and what is the object in a transitive clause (see the chapter by Malchukov and de Swart, this volume, for extensive discussion).

Aissen (1999, 2003) captures this idea of markedness in the interaction between on the one hand iconicity constraints such as 'Avoid unmarked animate objects', or to put it differently, 'Case mark animate objects', and on the other hand an economy constraint that penalizes morphological case marking. In a canonical transitive construction, the object is lower than the subject in animacy/definiteness, and thus when the object is animate/definite it is *marked* (for an object) which means it should be (case-)marked. For example, in Turkish specific objects are case-marked with accusative case, while non-specific objects remain without case (Enç 1991):

(15) *Ali bir kitab-ı aldı.*
Ali one book-ACC bought
'Ali bought a certain book.'

(16) *Ali bir kitap aldı.*
Ali one book bought
'Ali bought some book or other.'

Aissen crucially combines markedness of form with markedness of meaning within her constraints. She does that by local conjunction of independently motivated constraints that appeal to form (a constraint 'Case mark' which is comparable to the Case Filter known from generative grammar, cf. Chomsky 1981) and to meaning (constraints such as 'Avoid animate/specific objects'). Thus, she uses constraints such as 'Avoid a specific object which is not case marked' (*Oj/Spec & $*Ø_C$) and which is universally ranked higher than 'Avoid a non-specific object which is not case marked' (*Oj/Nspec & $*Ø_C$). By inserting the economy constraint 'Avoid case marking' ($*Struc_C$) in between these two constraints, Aissen accounts for the Turkish pattern in which specific indefinite objects are case-marked while non-specific indefinite objects are not. The ranking Aissen proposes for Turkish differential object marking is therefore 'Case mark a specific indefinite object' ≫ 'Avoid case marking' ≫ 'Case mark a non-specific indefinite object' (*Oj/Spec & $*Ø_C$ ≫ $*Struc_C$ ≫ *Oj/Nspec & $*Ø_C$). The optimization procedures for an input with a specific and a non-specific object are illustrated in the (simplified) tableaux 6.3 and 6.4.

Tableau 6.3. Derivation of a case-marked object in Turkish

Input: 'Ali bought a certain book.'	*Oj/Spec & $*Ø_C$	$*Struc_C$	*Oj/Nspec & $*Ø_C$
☞ Ali bir kitab-ı aldı		*	
Ali bir kitap aldı.	*!		

Tableau 6.4. Derivation of a case-marked object in Turkish

Input: 'Ali bought some book or other.'	*Oj/Spec & $*Ø_C$	$*Struc_C$	*Oj/Nspec & $*Ø_C$
Ali bir kitab-ı aldı		*!	
☞ Ali bir kitap aldı.			*

Aissen's approach thus elegantly accounts for cross-linguistic patterns in which accusative case is assigned to objects which are high in prominence (and thus less typical objects) and to subjects which are low in prominence (and thus less typical subjects). However, a problem for Aissen's analysis of differential case marking, pointed out by de Swart (2003) and Woolford (2001, 2008), is that it predicts differential subject marking to mirror differential object marking in the sense that case is always assigned to subjects which are low in prominence (and thus less typical subjects). This prediction is definitely borne out in some languages, such as for example Korean, where high-prominence (animate and indefinite) subjects often drop their case marking in colloquial speech (cf. Lee 2008). However, we find more examples of differential subject marking where in fact the subjects high in prominence rather than those low in prominence bear (ergative) case (cf. Malchukov and de Swart, this volume). These are subjects which are not less typical, and they are certainly not more object-like than low-prominence subjects. In these instantiations of differential subject marking, Legendre et al.'s (1993) constraint saying 'High-prominence arguments receive subject case (C_1)- marking' seems to be satisfied instead of Aissen's constraint saying 'Case mark the low-prominence (less typical) subject'.

Thus, in the domain of differential subject marking, Legendre et al.'s constraint, reformulated as 'High-prominence subjects receive ergative case' is in conflict with Aissen's constraint, rephrased as 'Low-prominence subjects receive ergative case'. On the other hand, with regard to differential object marking, there is no such conflict, and only one relevant constraint, namely 'High-prominence objects receive accusative case'. Now we can account for the cross-linguistic asymmetries that are found between differential subject marking and differential object marking. While the constraints boil down to one type of differential object marking (that is, marking the high-prominence object), in the case of differential subject marking we encounter different types of differential subject marking across languages, in accordance with the existence of the two conflicting constraints. Malchukov and de Swart (this volume) relate constraints such as used by Legendre et al. to the indexing function of case (i.e. the morphological realization of certain properties of the arguments), while the constraints by Aissen are instantiations of the distinguishing function of case (i.e. the morphological marking that is used to distinguish between subjects and objects). Interestingly, taking into account certain economy considerations as well, a correlation can be observed that explains why passivization is applied when there is need to encode a subject demotion in a nominative–accusative language, while antipassivization applies when an object demotion must be encoded in an ergative language (Malchukov 2006a).

Recall that the relation between prominence and voice was also accounted for by Legendre et al. (1993) who argue that passives occur when the input is *aP* (with a demoted subject), while antipassives are the result of an *Ap* input (with a demoted object). However, Legendre et al. do not account for the fact that passives are

found more often in nominative–accusative languages, while antipassives are found more often in ergative languages. In Malchukov's approach this is straightforwardly explained by the interaction between two conflicting constraints. The constraint that is of utmost importance in this conflict is dubbed PAIP, which is an abbreviation of 'Primary Actant Immunity Principle' (see also Malchukov and de Swart, this volume). PAIP penalizes case marking of an (otherwise) unmarked argument. Therefore, PAIP penalizes morphological case marking of the absolutive argument in ergative languages and of the nominative argument in nominative–accusative languages. PAIP thus penalizes 'marking the unmarked'. As has been argued by Malchukov (2006a), the potential conflict between case marking high-prominence arguments and PAIP can explain the striking fact that differential object marking is normally found in nominative–accusative languages, while differential subject marking is usually found in ergative–absolutive languages.

That is, in nominative–accusative languages, where the subject of a transitive sentence is the unmarked argument, differential object marking does not violate PAIP which is satisfied by the nominative subject. But in ergative languages, where the object of a transitive verb is the unmarked argument, object marking would induce a violation of PAIP. In those languages, we see that a low-prominence object leads to the use of an antipassive construction, while a strong object remains in the unmarked (absolutive) case. Thus, both the constraint that states that high-prominence arguments should be case-marked and PAIP can be satisfied in case of differential object marking in nominative–accusative languages and differential subject marking in ergative languages. But when the two constraints are in conflict, we commonly see that a voice alternation is a way to resolve the conflict (Malchukov 2006a). This is illustrated in the Table 6.1.

That is, passive constructions in nominative–accusative languages and antipassive constructions in ergative–absolutive languages can be analysed as optimal expressions of inputs with a demoted subject and object respectively, thus satisfying PAIP in the two types of languages. This constraint, however, is not violated by marking the subject in ergative–absolutive constructions (where the absolutive argument remains unmarked), nor by marking the object in nominative–accusative languages (where the nominative argument remains unmarked). Thus, we predict

Table 6.1. General distribution of differential case marking and voice alternations in different types of languages

Input	Nom–acc language	Erg–abs language
A P/p	DOM	Active/Antipassive
A/a P	Active/Passive	DSM

DOM to be possible in nominative–accusative languages and DSM in ergative–absolutive languages.

6.5 Bidirectional OT

So far, we have discussed examples of OT-syntax, related to the question what the optimal case marking is, given a semantic input. OT-semantic takes the other perspective (Hendriks and de Hoop 2001). That is, given a syntactic input with morphological case, what is its optimal interpretation? In OT one can proceed from the assumption that *in principle* there is total freedom in interpretation and generation. That is, for each utterance a hearer has to choose an interpretation from among an in principle infinite number of interpretations (the candidate set of interpretations). And similarly, a speaker has to choose an expression from among an in principle infinite number of structures (the candidate set of expressions). The question is how language users make these choices. Since Blutner (2000) the view has been established that the process of optimization is a bidirectional process (see also Jäger and Zeevat 2002 for a bidirectional version of Aissen's harmonic alignment framework). Crucially, speakers determine the optimal form by also taking into account the hearer's perspective who has to arrive at the intended meaning, while hearers determine the optimal interpretation by also taking into account the speaker's alternatives to express the intended meaning.

Typological analyses, especially those dealing with 'competing motivations' (Du Bois 1985), are perfectly compatible with OT syntax, as shown by the formal OT treatments briefly discussed in this chapter. Recently, however, also the use of bidirectional OT has been discovered to account for typological generalizations in the domain of case marking, in particular de Hoop and Malchukov (2006, 2007). Reconsider for example the pattern of fluid differential subject marking that certain intransitive verbs in Hindi show:

(17) Raam-ne jorse cillaayaa
Raam-ERG loudly shouted
'Raam shouted loudly (volitionally)'

(18) Raam jorse cillaayaa
Raam loudly shouted
'Raam screamed loudly'

The majority of intransitive verbs in Hindi do not allow for a subject marked with ergative case. This can be accounted for by a specific economy constraint that penalizes ergative case marking (Woolford 2001). On the basis of the potential

conflict between this constraint and a constraint which requires volitional subjects to be case-marked, we expect that ergative case marking of volitional subjects is obligatory when the latter constraint is ranked higher than the economy constraint, and that it is prohibited when the ranking is the other way around.

Hence, the majority of intransitive verbs in Hindi behave in accordance with the economy constraint ranked high. Thus, their subjects are in the nominative case, whether they are volitional or not. However, a small class of intransitive verbs, as illustrated in (17) and (18), would suggest the reverse ranking of the two constraints, leading to the ergative case marking of the volitional subjects. However, plain reranking of constraints within one language is not possible in Optimality Theory. If we allowed for it, OT would lose its explanatory power. One way to solve the problem pointed out above without reranking the constraints is to let another more specific constraint interfere with the two constraints. This third constraint would apply only in the context of the small set of verbs such as 'to shout' in Hindi, and not in the context of other intransitive verbs. But such an analysis would miss an important cross-linguistic generalization, as pointed out by de Hoop and Malchukov (2007). The tendency is that when a verb allows for two meanings of its subject, a volitional and a non-volitional meaning, and when at the same time it can have two different case markers on its subject, ergative and non-ergative case, then the ergative case is used for the volitional meaning, and the non-ergative case is used for the non-volitional meaning.

Whenever we encounter such a pattern in language where *in the same context* two forms are available as well as two meanings, we may find a mapping of forms and meanings that is in accordance with a principle of markedness that states that the marked form is used to express the marked meaning (Horn 1984). Blutner (2000) derives this markedness principle by means of bidirectional optimization (from form to meaning and from meaning to form). As argued by de Hoop and Narasimhan (2008) and de Hoop and Malchukov (2007), a bidirectional OT account can derive the marked combination of ergative form and volitional meaning in the context of a small set of verbs in Hindi, while keeping the original ranking of the relevant constraints intact. This is shown in the bidirectional OT tableau 6.5.

Tableau 6.5. Case on intransitive subject of 'shout' in Hindi

Subject of intransitive verb such as 'shout' in Hindi	*ERG	VOL→ERG
☞ –ERGATIVE, –VOLITIONAL		
–ERGATIVE, + VOLITIONAL		*
+ ERGATIVE, –VOLITIONAL	*	
☞ + ERGATIVE, + VOLITIONAL	*	

978–0–19–920647–6 07-Spencer-ch06 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 101 of 101 May 16, 2008 18:57

While the unmarked form (nominative case) is used for the non-volitional meaning (thus violating none of the constraints), the marked (ergative) form is used for the volitional meaning. The second winning pair differs in both form and meaning from the first pair and can therefore be bidirectionally optimal as well.

6.6 Conclusions

In this chapter I have discussed the rise of Optimality Theory in the domain of case marking. I argued that the very first application of OT in the domain of syntax by Legendre et al. (1993) has not lost much of its attractiveness yet, as it can account for voice and case alternations driven by prominence distinctions in the input. However, it needs to be extended to cover also patterns of differential case marking that occur independently of voice alternations. These case alternations are very often governed by prominence distinctions in the input as well, as recognized by Aissen (1999, 2003). I have demonstrated that the constraints proposed by Legendre et al. on the one hand and Aissen on the other, can reinforce each other in the domain of object marking, while they are in conflict in the domain of subject marking. This conflict is reflected in the typological data, as pointed out by de Hoop and Malchukov (2006, 2007). Finally, I have briefly discussed bidirectional OT as a recent development in OT that seems to be a new and promising way to deal with typological data, especially for categories that allow for alternations at the syntax–semantics interface, such as case.

Acknowledgement

I gratefully acknowledge the financial support of the Netherlands Organisation for Scientific Research (NWO) for the research reported here (grants no. 220-70-003 'Case cross-linguistically', DN 30-609 'Incremental interpretation of case and prominence', and 360-70-222 'Animacy').

CHAPTER 7

CASE IN ROLE AND REFERENCE GRAMMAR

ROBERT D. VAN VALIN, JR.

7.1 Introduction[1]

This paper will explicate the theory of case assignment in Role and Reference Grammar [RRG](Van Valin 2005). RRG is a monostratal, non-derivational theory which posits a direct linking between the syntactic and semantic representations of a sentence; discourse-pragmatics may play a role in this linking. Case marking in RRG is handled rather differently than in other theories: first, it is not assigned on the basis of grammatical relations; second, it is not assigned on the basis of phrase structure configurations; and third, it may directly reflect the semantic or pragmatic status of an argument, in some languages.

The discussion will begin with a summary of the linking system in RRG, concentrating on those aspects of it that are relevant to case marking. Then the basics of nominative–accusative and ergative–absolutive case assignment will be presented. Then the treatment of other significant argument-marking cases, i.e. the dative and the instrumental, will be given. There are complications affecting these case assignment rules, and the two most important will be explored. The final section

[1] I would like to thank Andrej Malchukov, Andrew Spencer, and an anonymous referee for comments. The writing of this paper was supported in part by a Research Award from the Alexander von Humboldt Stiftung.

978–0–19–920647–6 08-Spencer-ch07 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 103 of 120 May 16, 2008 18:57

leads to a discussion of languages in which case assignment is strongly affected by discourse-pragmatics.

7.2 Linking syntax and semantics in Role and Reference Grammar

RRG posits only a single syntactic representation for a sentence, which is the overt form of the sentence; there are no underlying syntactic structures, transformational rules, or derivations. This syntactic representation is related to the semantic representation of the sentence by a set of rules called the 'linking algorithm'.

The syntactic representation is known as the 'layered structure of the clause' and consists of two projections: the 'constituent projection' and the 'operator projection', which will not be discussed here. The constituent projection consists of the 'nucleus' of the clause, containing the predicate, the 'core' of the clause, containing the nucleus and the arguments of the predicate, and the 'periphery' of the clause, housing the adjuncts modifying the core. The structure of a simple sentence in English, constituent projection only, is exemplified in Figure 7.2.

A couple of notes are in order. NPs headed by common nouns and adjunct PPs have a layered structure analogous to that of clauses; NPs headed by proper nouns and pronouns lack a layered structure, as they do not take operators. The PP headed by *to* does not have a layered structure, because *to* is non-predicative, i.e. it does not license its object, *Chris*, which is an argument of the verb *present*, which can be seen clearly in the alternative form *Sandy presented Chris with the flowers.*

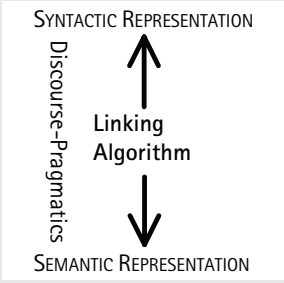


Figure 7.1. The organization of Role and Reference Grammar

Au: Please provide citation for figure 7.1 in text.

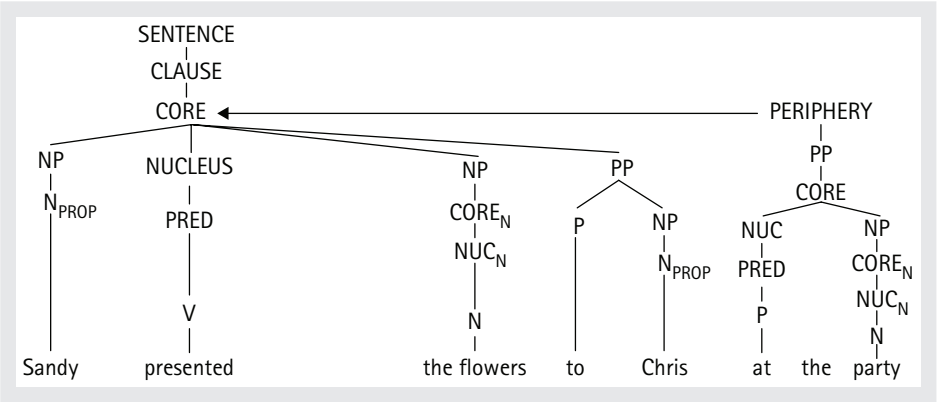


Figure 7.2. The layered structure of a simple English sentence

Syntactic structures are stored as syntactic templates in the syntactic inventory in the grammar. Syntactic templates are language-specific syntactic forms which are composed of the universal components of the layered structure of the clause. There are principles which determine the selection of syntactic templates for semantics-to-syntax linking; the default principle is that the core template must have as many argument slots as there are arguments in the semantic representation of the core.

The semantic representation of a sentence is based on the decompositional representation of the predicate in the nucleus. The decompositional system is based on the *Aktionsart* distinctions originally proposed in Vendler (1967), with some extensions. A subset of the classes is given in (1), with example sentences involving each type plus its causative counterpart given in (2).

(1) a. States: *be sick*, *be tall*, *be dead*, *love*, *know*, *believe*, *have*
b. Achievements: *pop*, *explode*, *collapse*, *shatter* (all intransitive)
c. Accomplishments: *melt*, *freeze*, *dry* (the intransitive versions); *learn*, *receive*
d. Activities: *march*, *walk*, *roll* (the intransitive versions); *swim*, *think*, *rain*, *read*, *eat*

(2)

a.	State:	The boy is afraid.
a′.	Causative state:	The dog frightens/scares the boy.
b.	Achievement:	The balloon popped.
b′.	Causative achievement:	The cat popped the balloon.
c.	Accomplishment:	The ice melted.
c′.	Causative accomplishment:	The hot water melted the ice.
d.	Activity:	The soldiers marched in the field.
d′.	Causative activity:	The sergeant marched the soldiers in the field.

Table 7.1. Lexical representations for some of the *Aktionsart* classes

Verb Class	Logical Structure (LS)
STATE	**predicate′** (x) or (x,y)
ACTIVITY	**do′** (x, [**predicate′** (x) or (x, y)])
ACHIEVEMENT	INGR **predicate′** (x) or (x,y), or INGR **do′** (x, [**predicate′** (x) or (x, y)])
ACCOMPLISHMENT	BECOME **predicate′** (x) or (x,y), or BECOME **do′** (x, [**predicate′** (x) or (x, y)])
CAUSATIVE	α CAUSE β, where α, β are LSs of any type

The decompositional system is adapted from that proposed in Dowty (1979); the sample classes are summarized in Table 7.1. Examples of some English sentences with their logical structures are given in (3).

(3) a. STATES

The window is shattered.	**shattered′** (window)
Fred is at the house.	**be-at′** (house, Fred)

b. ACTIVITIES

The children cried.	**do′** (children, [**cry′** (children)])
Carl ate snails.	**do′** (Carl, [**eat′** (Carl, snails)])

c. ACHIEVEMENTS

The window shattered.	INGR **shattered′** (window)
The balloon popped.	INGR **popped′** (balloon)

d. ACCOMPLISHMENTS

The snow melted.	BECOME **melted′** (snow)
Mary learned French.	BECOME **know′** (Mary, French)

e. CAUSATIVES

The dog scared the boy.	[**do′** (dog, Ø)] CAUSE [**feel′** (boy, [**afraid′**])]
Max broke the window.	[**do′** (Max, Ø)] CAUSE [BECOME **broken′** (window)]
The cat popped the balloon	[**do′** (cat, Ø)] CAUSE [INGR **popped′** (balloon)]
Felix bounced the ball.	[**do′** (Felix, Ø)] CAUSE [**do′** (ball, [**bounce′** (ball)])]

A key component of the RRG linking system is the system of semantic roles. There are two types. The first is thematic relations like effector, agent, experiencer, theme, and patient, which are defined in terms of argument positions in the decompositional representations. The second are semantic macroroles.[2] There are two

[2] See Van Valin (1999, 2004) for detailed discussion of semantic macroroles.

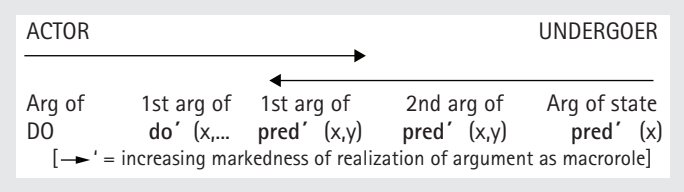


Figure 7.3. The Actor–Undergoer Hierarchy

semantic macroroles, actor and undergoer, which are the two primary arguments of a transitive predication; the single argument of an intransitive predicate can be either one, depending upon the semantics of the verb. This is illustrated in (4).

(4) a. Kim [Actor] ate the bagel [Undergoer].
b. The bagel [Undergoer] was eaten by Kim [Actor].
c. Chris [Actor] jogged in the park.
d. Pat [Undergoer] fell asleep in class.

Actor and undergoer are called 'macroroles' because each of them subsumes a number of thematic relations; this is illustrated in (5).

(5)

	Actor	Undergoer
a. The farmer [A] killed the duckling [U].	Agent	Patient
b. The rock [A] broke the window [U].	Instrument	Patient
c. The lawyer [A] received the summons [U].	Recipient	Theme
d. Many tourists [A] saw the accident [U].	Experiencer	Stimulus
e. Sally [A] presented Bill [U] with the award.	Agent	Recipient
e′. Sally [A] presented the award [U] to Bill.	Agent	Theme
f. The mugger [A] robbed Sam [U] of $50.	Agent	Source
f′. The pickpocket [A] stole $50 [U] from Sam.	Agent	Theme
g. The clown [A] amused the child [U].	Agent	Experiencer

The relationship between argument positions in logical structure and actor and undergoer selection is expressed in the Actor–Undergoer Hierarchy; it is given in Figure 7.3.[3] This hierarchy states that the leftmost argument in the logical structure will be the actor and the rightmost the undergoer. While the actor selection principle is absolute and invariable across languages, there is variation with respect to undergoer selection; namely, with some verbs in some languages, it is possible to select a higher-ranked argument as undergoer. This is exemplified in the English dative shift and transfer alternations in (6) (where 'NMR' means 'non-macrorole argument').

[3] Certain complications arise in what are called 'primary-object languages'(Dryer 1986); see Guerrero and Van Valin (2004) and Van Valin (2005) for detailed discussion.

(6) a. Sally gave the flowers [U] to Kim [NMR].
a′. Sally gave Kim [U] the flowers [NMR].
b. Sally presented the flowers [U] to Kim [NMR].
b′. Sally presented Kim [U] with the flowers [NMR].
c. [**do**′ (Sally, Ø)] CAUSE [BECOME **have**′ (Kim, flowers)]

In an active voice English core, the undergoer is the direct NP that immediately follows the nucleus. In (6a,b) the lowest-ranking argument in the logical structure in (6c) is selected as undergoer; this is the default or unmarked selection. In (6a′,b′), on the other hand, the second lowest-ranking argument has been selected as undergoer, yielding a marked selection.

Transitivity is defined in terms of the number of macroroles that a verb takes, not the number of syntactic arguments. Accordingly, a verb with no macrorole arguments, e.g. *snow*, is M-atransitive, one with one macrorole argument is M-intransitive, e.g. *die*, and one with two macrorole arguments is M-transitive, e.g. *crush*, *put*. The default principle governing the assignment of macroroles is given in (7).

(7) Default Macrorole Assignment Principles
a. Number: the number of macroroles a verb takes is less than or equal to the number of arguments in its logical structure.
1. If a verb has two or more arguments in its LS, it will take two macroroles.
2. If a verb has one argument in its LS, it will take one macrorole.
b. Nature: for verbs which take one macrorole:
1. If the verb has an activity predicate in its LS, the macrorole is actor.
2. If the verb has no activity predicate in its LS, the macrorole is undergoer.

When the number of macroroles a verb takes does not follow from (7a), then the number of macroroles must be specified in the lexical entry of the verb by a feature MRa, where a is the number of macroroles. When the macrorole number is one, the principle in (7b) specifies which one it is.

Subject selection (or in RRG terms, 'privileged syntactic argument'[PSA] selection), is based on the hierarchy in (8) and the principles in (9).

(8) Privileged syntactic argument [subject]selection hierarchy:

arg of DO > 1st arg of **do**′ > 1st arg of **pred**′ (x, y) > 2nd arg of **pred**′ (x, y) > arg of **pred**′ (x)

(9) Privileged syntactic argument ['subject'] selection principles:
a. Accusative constructions: Highest-ranking direct core argument in terms of (8)
b. Ergative constructions: Lowest-ranking direct core argument in terms of (8)

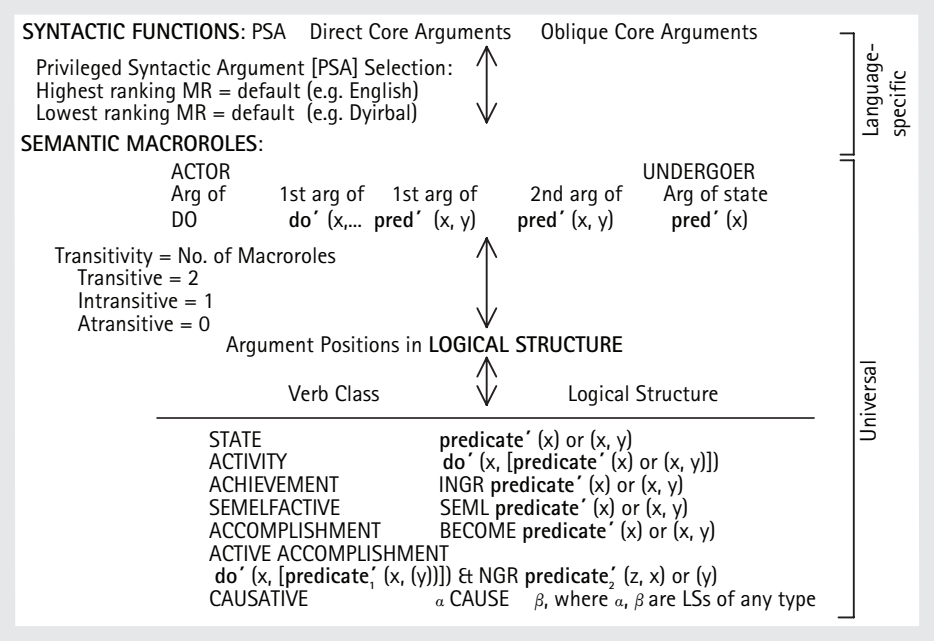


Figure 7.4. Summary of RRG linking system

c. Restrictions on PSA in terms of macrorole status:
   1. Languages in which only macrorole arguments can be PSA: German, Croatian, . . .
   2. Languages in which non-macrorole arguments can be PSA: Icelandic, Belhare, . . .

The protoypical non-macrorole PSA is a dative NP which has behavioural 'subject' properties; see Van Valin and LaPolla (1997: §7.3.1.1), Bickel (2006) for detailed discussion. In an accusative language like English, the default choice for subject is the highest-ranking macrorole in terms of (8), which would be the actor. It is possible to override this in a passive construction, in which the undergoer functions as subject (cf. (4b)). The components of the RRG linking system are summarized in Figure 7.4.

In Figure 7.1, discourse-pragmatics is mentioned, and it plays a significant role in the linking algorithm, one which varies in important ways across languages. Its primary manifestation in RRG is in terms of the notion of 'focus or information structure'. Focus structure is the grammatical system which serves to indicate the scope of the assertion in an utterance in contrast to the pragmatic presupposition (Lambrecht 1994), and it is vital to the RRG analysis of many grammatical phenomena, e.g. constraints on pronominalization (Van Valin and LaPolla 1997: §5.6), the interpretation of quantifier scope (ibid.: §5.5; Van Valin 2005: §3.6), the origin of VPs in languages that have them (Van Valin 2005: §3.5), and extraction restrictions (Van Valin 1995, 1998; Van Valin and LaPolla 1997: §9.5; Van Valin 2005:

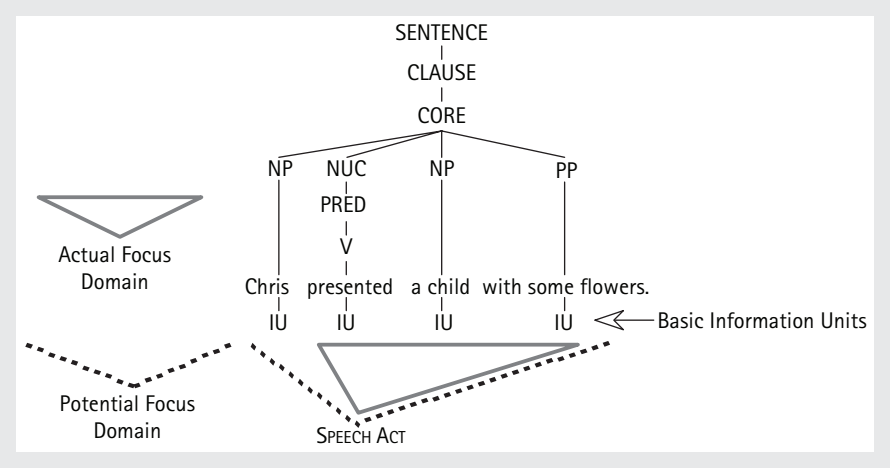


Figure 7.5. Predicate Focus Construction in English

§7.6). An innovation in RRG is the distinction between the 'potential focus domain' [PFD] i.e. the syntactic domain in the sentence where focus may fall, and the 'actual focus domain', i.e. the part that is focused in a particular utterance. Languages vary in terms of how the PFD is restricted, both in simple sentences and in complex sentences, and this variation underlies important grammatical differences across languages (Van Valin and LaPolla 1997: §5.3, 7.6). The focus structure of an utterance is represented in a distinct projection of the clause from the operator and constituent projections; this is exemplified in Figure 7.5 for a predicate focus construction in English. 'Predicate focus' is Lambrecht's (1994) term for the traditional 'topic-comment' structure with a topical subject and a focal predicate; 'IU' stands for 'basic information unit'.

Another important focus structure type is what Lambrecht calls 'sentence focus', a topic-less sentence in which the entire sentence is in the actual focus domain. It is captured in the following question–answer pair.

(10) a. What happened?
b. My CAR broke down.

The question does not establish a topic, and as a result the entire clause in (10b) is in the actual focus domain.

The linking between syntax and semantics is subject to a general constraint called the 'Completeness Constraint'; it is given in (11).

(11) Completeness Constraint
All of the arguments explicitly specified in the semantic representation of a sentence must be realized syntactically in the sentence, and all of the referring expressions in the syntactic representation of a sentence must be linked to an argument position in a logical structure in the semantic representation of the sentence.

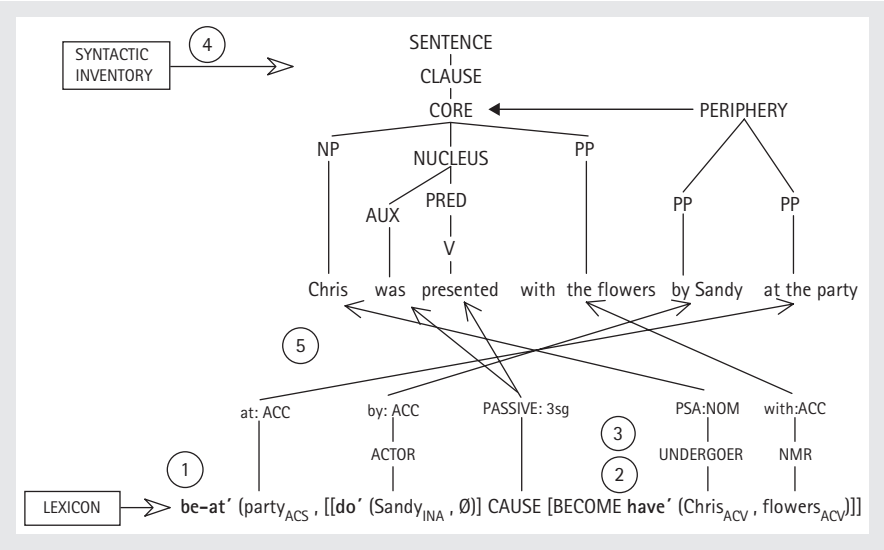


Figure 7.6. Linking from semantics to syntax in a simple sentence in English

A simple example from English illustrating the operation of the semantics-to-syntax linking algorithm is given in Figure 7.6. The numbers refer to the general steps of the algorithm: (1) constructing the semantic representation of the sentence; (2) assigning actor and undergoer; (3) determining PSA selection, case and adposition assignment, and agreement; (4) selecting the appropriate syntactic template from the syntactic inventory; and (5) linking the elements from the semantic representation into the appropriate positions in the syntactic representation. The subscripts 'ACV' and 'ACS' stand for 'activated' and 'accessible', and they refer to different cognitive statuses that a referent of the element may have; cf. Lambrecht (1994). Because this sentence is a passive, the undergoer appears as the 'subject', with the actor appearing in a peripheral PP. These language-specific details would be represented in the constructional schema for the English passive. See Van Valin (2005) for detailed discussion and explication of all of these points.

7.3 Case assignment

Case assignment in RRG cannot be based on either grammatical relations (functions) or phrase-structure configurations. The theory posits only a single grammatical relation, privileged syntactic argument; there is nothing corresponding to direct and indirect object. Obviously, the three primary cases in a language,

i.e. nominative, accusative, and dative, or ergative, absolutive, and dative, cannot be accounted for in terms of a single grammatical function. Due to the monostratal nature of the theory, phrase structure configurations cannot be the basis of case assignment, either. In languages with flexible word order, such as Russian, there is no particular position in the clause associated with any particular case, and there is no possibility in a theory like RRG of setting up abstract underlying forms in which case and position correlate. Hence case assignment cannot be based on phrase structure.

The case-marking rules cover regular (non-idiosyncratic) case marking by verbs and are given in (12) and (13); they apply to direct core arguments only.

(12) Case assignment rules for accusative constructions:
   a. Assign nominative case to the highest-ranking macrorole argument (in terms of (8)).
   b. Assign accusative case to the other macrorole argument.

(13) Case assignment rules for ergative constructions:
   a. Assign absolutive case to the lowest-ranking macrorole argument (in terms of (8)).
   b. Assign ergative case to the other macrorole argument.

The hierarchy in (8) refers to direct core arguments, not to macroroles, but they are related: the actor is the highest-ranking direct core argument, since by the actor–undergoer hierarchy the highest-ranking direct core argument is selected as actor, and conversely the undergoer will always be the lower-ranking of the two macrorole arguments, in terms of (8).

The application of the rules in (12) can be seen in the following Russian examples.

(14) a. *Učitel'nic-a* *pro-čita-l-a* *knig-u.*
   teacher-F.SG.NOM PFV-read-PAST-F.SG book-F.SG.ACC
   'The teacher read the book.'
   a′. **do**′ (učitel'nic-, [**read**′ (učitel'nic-, knig-)]) & INGR **consumed**′ (knig-)
   b. *Učitel'nic-a* *govori-l-a.*
   teacher-F.SG.NOM speak-PAST-F.SG
   'The teacher was speaking.'
   b′. [**do**′ (učitel'nic-, [**speak**′ (učitel'nic-)]
   c. *Ženščin-a* *umer-l-a.*
   woman-F.SG.NOM die-PAST-F.SG
   'The woman died.'
   c′. BECOME **dead'** (ženščin-)

The first example contains the M-transitive verb *pročitat'* 'read [completely]'; the actor *učitel'nic-* 'teacher' is the highest-ranking macrorole and therefore receives nominative case, while the undergoer *knig-* 'book' is the other macrorole and

978–0–19–920647–6 08-Spencer-ch07 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 112 of 120 May 16, 2008 18:57

therefore receives accusative case. The last two examples contain M-intransitive verbs, one of which takes an actor (*govorit'* 'speak' in (14b)) and the other an undergoer (*umeret'* 'die' in (c)). Since the single macrorole, regardless of type, is the highest-ranking by virtue of being the only one, it receives nominative case.

The application of the ergative case assignment rules in (13) is illustrated in the Dyirbal examples in (15) from Dixon (1972); these rules are the basic ones for an ergative system, and they do not deal with the complexities that arise in split-ergative systems (see Van Valin and LaPolla 1997: §7.3.1.2). ('NM' stands for 'noun-class marker'.)

(15) a. *Balan dyugumbil-Ø bani-ɲu.*
NM-ABS woman-ABS come-TNS
'The woman is coming.'
a′. **do**′ (dyugumbil-, [**come**′ (dyugumbil-)])
b. *Balan dyugumbil-Ø baŋgul* yaŗa-ŋgu buŗa-n.
NM-ABS woman-ABS NM-ERG man-ERG see-TNS
'The man sees the woman.'
b′. **see**′ (yaŗa-, dyugumbil-)

The first example contains an M-intransitive verb, *bani-* 'come'; it takes an actor argument. It is the only macrorole in the core, hence lowest-ranking, and therefore receives absolutive case, following (13a). In (b) with an M-transitive verb *buŗa-* 'see', *dyugumbil-* 'woman' is the undergoer and the lowest-ranking macrorole, while *yaŗa-* 'man' is the actor and the other macrorole. Hence *dyugumbil-* appears in the absolutive case and *yaŗa-* in the ergative case.

The rules for assigning dative and instrumental case are the same for both ergative and accusative systems. They are given in (16).

(16) a. Assign instrumental case to non-MR *b* argument if, given two arguments, *a* and *b*, in a logical structure, with (1) both as possible candidates for a particular macrorole and (2) *a* is equal or higher (to the left of *b*) on the AUH, *b* is not selected as that macrorole.
b. Assign dative case to non-macrorole arguments (default).

The idea that dative is the default case for non-macrorole core arguments comes from Silverstein (1981). Instrumental is not a default case and has specific conditions on its application. The assignment of instrumental case in Croatian and Dyirbal is illustrated below. (The Croatian example is from Ranko Matasović, personal communication; the Dyirbal example is from Dixon 1972.)

978–0–19–920647–6 08-Spencer-ch07 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 113 of 120 May 16, 2008 18:57

(17) a. *Žena-Ø je otključa-l-a vrata* Croatian
woman-F.SG.NOM be.3SG unlock-PAST-F door.N.PL.ACC
ključ-em.
key-M.SG.INS
'The woman unlocked the door with the key.'
b. [**do'** (žena-, Ø)] CAUSE [[**do'**(ključ-, Ø)] CAUSE [BECOME NOT **locked'** (vrat-)]]

(18) a. *Bala yugu-Ø baŋgul yaŗa-ŋgu baŋgu bari-ŋgu* Dyirbal
NM.ABS tree-ABS NM.ERG man-ERG NM.INS axe-INS
nudi-n.
cut-NFUT
'The man cut the tree with an axe.'
b. [**do'** (yaŗa-,Ø)] CAUSE [[**do'** (bari-, Ø)] CAUSE [BECOME **cut'** (yugu-)]]

In both of these logical structures, there are two effector arguments: *žena-* 'woman' and *ključ-* 'key' in (17b) and *yaŗa-* 'man' and *bari-* 'axe' in (18b); both are potential actors. The first effector in each logical structure is chosen as actor, leaving the second effector as a non-macrorole argument. This satisfies the conditions specified in (16a), and therefore instrumental case is assigned to the non-actor effectors in both sentences.

The application of the dative rule can be exemplified with sentences from Russian and Dyirbal.

(19) a. *Učitel'nic-a da-l-a knig-u ženščin-e.*
teacher-F.SG.NOM give-PAST-F.SG book-F.SG.ACC woman-F.SG.DAT
'The teacher gave a book to the woman.'
b. [**do'** (učitel'nic-, Ø)] CAUSE [BECOME **have'** (ženščin-, knig-)]

The linking of actor and undergoer is the same as in (14a): *učitel'nic-* 'teacher' is the highest-ranking argument and therefore the actor, while *knig-* 'book' is the lowest-ranking argument and therefore the undergoer. They receive the same cases as in (14a). The third core argument, *ženščin-* 'woman', is a non-macrorole core argument and because the conditions for instrumental case do not apply, it receives dative case.

(20) a. *Balam mirañ-Ø baŋgun dyugumbi-ŗu wuga-n bagul yaŗa-gu.*
NM.ABS beans-ABS NM.ERG woman-ERG give-TNS NM.DAT man-DAT
'The woman gave beans to the man.'
b. [**do**′ (dyugumbil-, Ø)] CAUSE [BECOME **have**′ (yaŗa-, mirañ-)]

In the Dyirbal example, the ditransitive verb *wuga-* 'give' heads the core; *dyugumbil-* is the actor and *miraŋ-* 'beans' is the undergoer, following the actor–undergoer

978–0–19–920647–6 08-Spencer-ch07 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 114 of 120 May 16, 2008 18:57

hierarchy, and *yaŋa-* 'man' is left as a non-macrorole core argument. Accordingly, *dyugumbil-* receives ergative case, *miraɲ-* absolutive case, and *yaŋa-* dative case.

The dative and instrumental rules interact in an interesting way with verbs which permit variable undergoer choice, and such verbs are found in both Croatian and Dyirbal. Examples illustrating this alternation are given in (21) from Croatian (Zovko 2000, 2001) and in (20a) and (22) from Dyirbal (Dixon 1972).

(21) *darovati* 'give as a gift' [**do**′ (unuk-, Ø)] CAUSE [BECOME **have**′ (bak-, cvijet-)]

a. *Unuc-i* *su* *bak-i* *darova-l-i*
grandson-M.PL.NOM be.3PL grandmother-F.SG.DAT give-PAST-PL
cvijeć-e.
flower-M.PL.ACC
'The grandsons gave flowers to [their] grandmother.'

b. *Unuc-i* *su* *bak-u* *darova-l-i*
grandson-M.PL.NOM be.3PL grandmother-F.SG.ACC give-PAST-PL
cvijeć-em.
flower-M.PL.INS
'The grandsons gave [their] grandmother flowers.'

(22) *Bayi* *yaŋa-Ø* *wuga-n* *baŋgun* *dyugumbi-ŋu* *baŋgum* *miraɲ-dyu.*
NM.ABS man-ABS give-TNS NM.ERG woman-ERG NM.INS beans-INS
'The woman gave the man beans.'

The Croatian verb *darovati* 'give as a gift' allows either the recipient or the theme to function as undergoer. In the default linking, illustrated in (21a), *cvijet-* 'flowers', the lowest-ranking argument, is selected as undergoer, and *bak-* 'grandmother', the non-macrorole core argument, receives dative case. In the other linking, shown in (21b), the recipient *bak-* 'grandmother' is chosen as undergoer, leaving *cvijet-* 'flowers' as a non-macrorole argument. This linking meets the conditions of the instrumental case rule: both *cvijet-* 'flowers' and *bak-* 'grandmother' are potential undergoers, and *bak-* 'grandmother' is to the left of *cvijet-* 'flowers' in the logical structure. Consequently, *cvijet-* 'flowers' is assigned instrumental case. The same contrast is found in the Dyirbal examples in (20a) and (22). The linking in the first sentence is described above. In (22), the recipient *yaŋa-* 'man' is chosen as the undergoer, leaving *miraɲ-* 'beans' a non-macrorole argument. Here too the instrumental rule can apply, because *yaŋa-* 'man' is higher on the actor–undergoer hierarchy than *miraɲ-* 'beans'; accordingly, *miraɲ-* 'beans' receives instrumental case. In these pairs of sentences, dative case on the non-macrorole argument correlates with the unmarked choice for undergoer in terms of the actor–undergoer hierarchy, while instrumental case on it correlates with a marked choice for undergoer, i.e. the undergoer is not the lowest-ranking argument in the logical structure.

Dative case in RRG is one of the regular cases; it is not semantic or idiosyncratic ('quirky'). When verbs take an 'unexpected' dative 'subject' or 'object', e.g. German

gefallen 'be pleasing to' (dative 'subject') or *helfen* 'help' (dative 'object'), this is the result of the verb being M-intransitive in terms of (7a). The logical structures for *gefallen* and *helfen* are **please**′ (x, y) [MR1] and **do**′ (x, [**help**′(x, y)]) [MR1], respectively. By the principles in (7b), the single macrorole with *gefallen* is undergoer, which is the *y* argument following the actor–undergoer hierarchy, while the single macrorole with *helfen* is actor, which would be the *x* argument. The remaining argument in each logical structure is a non-macrorole direct core argument. With *gefallen*, the undergoer is the highest-ranking macrorole and receives nominative case, following (12a), and with *helfen* the actor is highest ranking and is assigned nominative case. The non-macrorole core arguments receive dative case, following (16b). The 'quirky' feature of these verbs is their unexpected M-intransitivity; their case marking is completely regular. See Van Valin and LaPolla (1997: §7.3.1.1) for detailed discussion.

7.4 Semantics and focus structure in case assignment

Case assignment is typically viewed as being syntactically or semantically motivated, but Korean gives evidence that focus structure can motivate case assignment in some instances. Korean exhibits the unusual phenomena of case spreading and case stacking; Park (1995) and Han (1999) present RRG analyses of them. An example of case spreading with an intransitive verb is given in (23), and examples of case spreading and case stacking with a transitive verb are given in (24); they are taken from Han (1999).

(23) a. *Thoyoil-ey kongcang-eyse pwul*-i na-ass-ta. Basic form
Saturday-LOC factory-LOC fire-NOM break.out-PAST-DECL
'Fire broke out in the factory on Saturday.'

b. *Thoyoil-i kongcang-i pwul-i* Case spreading
Saturday-NOM factory-NOM fire-NOM
na-ass-ta.
break.out-PAST-DECL
'Fire broke out in the factory on Saturday.'

(24) a. *Chelswu-ka Yenghi-eykey kkot-ul cwu-ess-ta.* Basic form
Chelswu-NOM Yenghi-DAT flower-ACC give-PAST-DECL
'Chelswu gave a flower to Yenghi.'

b. *Chelswu-ka Yenghi-lul kkot-ul cwu-ess-ta.* Case spreading
Chelswu-NOM Yenghi-ACC flower-ACC give-PAST-DECL
'Chelswu gave Yenghi a flower.'

978–0–19–920647–6 08-Spencer-ch07 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 116 of 120 May 16, 2008 18:57

c. *Chelswu-ka* *Yenghi-eykey-lul* *kkot-ul* Case stacking
Chelswu-NOM Yenghi-DAT-ACC flower-ACC
cwu-ess-ta.
give-PAST-DECL
'Chelswu gave a flower to Yenghi.'

Case spreading occurs when the nominative case occurs on more than just the PSA in a clause with an intransitive verb, as in (23b) or when the accusative appears on more than just the undergoer in a clause with a transitive verb, as in (24b). All three NPs in (23b) are in the nominative case, which is rather unusual from a cross-linguistic perspective and raises important questions about the nature of Korean case marking. In (24b) there are two accusative NPs, the undergoer and the recipient, whereas in the basic form in (a), the undergoer is in the accusative case and the recipient is in the dative, following the usual case marking rules for accusative languages in (12) and the dative rule in (16b). In (24c), the recipient NP, *Yenghi,* carries two cases, dative and accusative; this is case stacking. In case spreading, one case replaces another, whereas in case stacking, the accusative case is added as a second case marker on certain NPs. Both case spreading and case stacking are very unusual cross-linguistically, and it is even more striking that both occur in the same language. In (24b) and (c) there are two accusative NPs; what is the difference between the two forms?

A clue to the function of the accusative case in these sentences can be found in the following examples involving verbs of motion, which do not normally take an accusative argument.

(25) a. *Chelswu-ka* *san-ey* *kan-ess-ta.* Basic form
Chelswu-NOM mountain-LOC go-PAST-DECL
'Chelswu went to(wards) the mountain.'

a′. *Chelswu-ka* *san-ey* *kan-ess-ciman,* *ku-nun* *tochankha-ci*
Chelswu-NOM mountain-LOC go-PAST-but he-TOP arrive-CLM
an-ass-ta.
NEG-PAST-DECL
'Chelswu went to(wards) the mountain, but he did not arrive.'

b. *Chelswu-ka* *san-lul* *kan-ess-ta.* Accusative form
Chelswu-NOM mountain-ACC go-PAST-DECL
'Chelswu went to the mountain.'

b′. **Chelswu-ka* *san-lul* *kan-ess-ciman,* *ku-nun* *tochankha-ci*
Chelswu-NOM mountain-ACC go-PAST-but he-TOP arrive-CLM
an-ass-ta.
NEG-PAST-DECL
'Chelswu went to the mountain, but he did not arrive.'

In (25b) the accusative case has replaced the locative case on *san* 'mountain', the goal of the motion verb *kan* 'go', and this signals an important semantic contrast. As

(25a′) shows, (25a) does not necessarily entail that Chelswu reached the mountain, whereas in the accusative version in (b), it is entailed that Chelswu reached the mountain, as the contradiction in (b′) shows. Thus (25b) must be interpreted as an active accomplishment, while (25a) can be interpreted as an activity. This difference in required telicity is signalled by the replacement of the locative case by the accusative in (25b). Is there a comparable contrast between (24a) and (24b)?

(26) a. *Chelswu-ka Yenghi-eykey kkot-ul cwu-ess-ciman*, Basic form
Chelswu-NOM Yenghi-DAT flower-ACC give-PAST-but
Yenghi-nun pat-ci an-ass-ta.
Yenghi-TOP have-CLM NEG-PAST-DECL
'Chelswu gave a flower to Yenghi, but Yenghi did not have it.'

b. **Chelswu-ka Yenghi-ul kkot-ul cwu-ess-ciman*, Case spreading
Chelswu-NOM Yenghi-ACC flower-ACC give-PAST-but
Yenghi-nun pat-ci an-ass-ta.
Yenghi-TOP have-CLM NEG-PAST-DECL
'Chelswu gave Yenghi a flower, but Yenghi did not have it.'

The contrast between the two sentences in (26) parallels that between (25a′) and (b′): (26a) does not necessarily entail that Yenghi received the flower, while (26b) does entail completion of the transfer. Here again the replacement of the case on the 'goal' argument by the accusative signals completion. It appears, then, that the 'spreading' of the accusative case is semantically motivated and signals that an action is telic and completed; it will henceforth be referred to as 'accusative substitution'.[4]

The nominative spreading in (23) has a different explanation. In the basic form in (23a), the undergoer, the highest-ranking macrorole, is assigned nominative case by (12a), and the other two NPs are adjuncts carrying variants of the locative case. In (23b), on the other hand, all three NPs are in the nominative case. The motivation for this can be found in the contexts in which the different forms are used, as shown in (27).

(27) a. *Mwusenil-i ilena-ess-ni*?
what-NOM happen-PAST-Q
'What happened?'

a′. *Thoyoil-i kongcang-i pwul-i na-ess-ta*
Saturday-NOM factory-NOM fire-NOM break.out-PAST-DECL
'Fire broke out in the factory on Saturday.'

b. *Thoyoil-ey mwusenil-i ilena-ess-ni?*
Saturday-on what-NOM happen-PAST-Q
'What happened on Saturday?'

[4] Given that verbs of motion do not normally take an accusative NP, the phenomenon in (25) is not really case 'spreading', but it is clearly related to the accusative 'spreading' in (24); consequently the term 'accusative substitution' will be used to refer to both phenomena.

978–0–19–920647–6 08-Spencer-ch07 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 118 of 120 May 16, 2008 18:57

b′. *Thoyoil-ey/*i kongcang-i pwul-i na-ess-ta*
Saturday-LOC/NOM factory-NOM fire-NOM break.out-PAST-DECL
'Fire broke out in the factory on Saturday.'

c. *Thoyoil-ey kongcang-eyse mwusenil-i ilena-ess-ni?*
Saturday-LOC factory-LOC what-NOM happen-PAST-Q
'What happened at the factory on Saturday?'

c′. *Thoyoil-ey/*i kongcang-eyse/*i pwul-i na-ess-ta*
Saturday-LOC/NOM factory-LOC/NOM fire-NOM break out-PAST-DECL
'Fire broke out in the factory on Saturday.'

The question in (27a) is the same one used in (10) to create the context for a sentence focus construction, and it is in this context that all three NPs appear in the nominative case. When the question establishes a topic, as in (27b) and (c), then the NP mentioned in the question cannot occur in the nominative case in the answer. Thus it would appear that the 'extra' nominative case signals that the NP is part of the actual focus domain; in other words, it indicates that an NP is focal.

The means of signalling the actual focus domain in clauses with transitive verbs is case stacking. The accusative case added to the dative NP *Yenghi-eykey* 'to Yenghi' in (24c) highlights that this NP is focal. Two pieces of evidence for this can be found in Korean questions. Wh-questions are narrow focus constructions, with the wh-word being the focused element. Consequently, it should be impossible to have case stacking in a wh-question, because this would lead to a conflict regarding which NP is the focus of the sentence. This is correct, as (28b) shows.

(28) a. *Chelswu-ka Yenghi-eykey mwuet-ul cwu-ess-ni?*
Chelswu-NOM Yenghi-DAT what-ACC give-PAST-Q
'What did Chelswu give to Yenghi?'

b. **Chelswu-ka Yenghi-eykey-lul mwuet-ul cwu-ess-ni?* Case stacking
Chelswu-NOM Yenghi-DAT-ACC what-ACC give-PAST-Q
'What did Chelswu give to Yenghi?'

A similar problem arises with negation. As has long been known (e.g. Russell 1905, Jackendoff 1972, Sgall et al. 1986), the scope of negation in a sentence correlates with focus; in other words, in a sentence with a negative, its scope is the actual focus domain. Consider the following question–answer pairs; small caps indicate focal stress.

(29) a. *Chelswu-ka Yenghi-eykey kkot-ul* CWU-ESS-NI?
Chelswu-NOM Yenghi-DAT flower-ACC give-PAST-Q
'Did Chelswu GIVE a flower to Yenghi?'

b. *Ani, Chelswu-ka Yenghi-eykey kkot-ul* Basic form
no Chelswu-NOM Yenghi-DAT flower-ACC
an-CWU-ESS-TA.
NEG-give-PAST-DECL
'No, Chelswu did not GIVE a flower to Yenghi.'

c. **Ani, Chelswu-ka Yenghi-eykey-lul kkot-ul* Case stacking
no Chelswu-NOM Yenghi-DAT-ACC flower-ACC
an-CWU-ESS-TA.
NEG-give-PAST-DECL
'No, Chelswu did not GIVE a flower to Yenghi.'

The question in (29a) has narrow focus on the verb, and in the reply in (b), the scope of the negative *an-* is just the verb. In (c), on the other hand, there is a conflict between the stacked accusative on *Yenghi-eykey* 'to Yenghi', which highlights that it is focal, and the prefixal negative plus focal stress on the verb, which indicates that it is the focused element in the clause. The result is an unacceptable response to (29a). If *Yenghi-eykey* had been the focus of the question and hence the focus of the response, then (29c) would be fine. The data in (28) and (29) support the analysis of stacked accusative case as a focus marker.

Thus, accusative case substitution and case stacking have different functions in Korean; they are not simply morphological variants of each other. This can be seen clearly in the contrast between (25b′) and (26b) with accusative substitution, on the one hand, and the comparable sentence with case stacking in (30) and (31).

(30) a. *Chelswu-ka san-ey-lul kan-ess-ta.* Case stacking
Chelswu-NOM mountain-LOC-ACC go-PAST-DECL
'Chelswu went to the mountain.'

b. *Chelswu-ka san-ey-lul kan-ess-ciman, ku-nun*
Chelswu-NOM mountain-LOC-ACC go-PAST-but he-TOP
tochankha-ci an-ass-ta.
arrive-CLM NEG-PAST-DECL
'Chelswu went to the mountain, but he did not arrive.'

(31) a. *Chelswu-ka Yenghi-eykey-lul kkot-ul* Case stacking
Chelswu-NOM Yenghi-DAT-ACC flower-ACC
cwu-ess.
give-PAST
'Chelswu gave a flower to Yenghi.'

b. *Chelswu-ka Yenghi-eykey-lul kkot-ul cwu-ess-ciman,*
Chelswu-NOM Yenghi-DAT-ACC flower-ACC give-PAST-but
Yenghi-nun pat-ci an-ass-ta.
Yenghi-TOP have-CLM NEG-PAST-DECL
'Chelswu gave a flower to Yenghi, but Yenghi did not have it.'

Adding the accusative case on top of the locative or the dative does not entail completion or telicity, as the acceptability of the (b) examples shows. Hence replacing the dative or locative by the accusative has different semantic consequences from stacking the accusative case on them.

978–0–19–920647–6 08-Spencer-ch07 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 120 of 120 May 16, 2008 18:57

The Korean case phenomena in this section have illustrated two important points: first, case assignment can be motivated by focus structure considerations, as in nominative spreading and accusative stacking, and second, it can be motivated by semantic concerns such as telicity, as in accusative substitution. For a language like Korean, case rules like the ones in (12) cover only the basic syntactic uses of the cases; additional rules would be required to handle the phenomena in this section, and Han (1999) formulates such rules.

7.5 Conclusion

RRG provides a rather different approach to the analysis of case. Case assignment is based neither on grammatical relations nor on phrase structure, as in many other theories, and it can be directly motivated semantically or pragmatically. Furthermore, dative and instrumental case, normally treated as semantic, lexical, or idiosyncratic, are treated as regular cases in RRG, which are assigned by general rules.

CHAPTER 8

CASE IN LOCALIST CASE GRAMMAR

JOHN M. ANDERSON

In the centuries immediately preceding the twentieth we can differentiate, simplifying somewhat, two main attitudes to the study of case. On the one hand, there was reasserted, particularly in the nineteenth century, a view of case as a purely **morphological** phenomenon (reverting to prevalent conceptions in the ancient world). These subtypes of case inflections came to be characterized by the functions they signalled, with a major division being made between 'syntactic cases' and 'local'. For Rumpel (1845, 1866), for instance, the nominative is the case of the 'subject' of the verb, the accusative that of the 'direct object', the dative that of the 'indirect object', and the genitive marks 'subjects' or 'objects' of nouns. However, other case inflections cannot be so interpreted – and indeed in many languages the inflections that are regarded as marking the 'direct object' and 'indirect object', and even the 'subject', also have other, typically 'local', functions.[1] These 'non-syntactic uses', and 'cases', can be given a 'local', spatial, interpretation.

[1] Even if we assume that accusative marks 'direct object' in Latin, as in (i), it also can signal a spatial goal, as in (ii):

(i) *Immodica īra gīgnit īnsāniam*
excessive anger it-causes madness:ACC ('Excessive anger causes madness')

(ii) *Missī lēgātī Athēnās sunt*
sent envoys: NOM Athens:ACC are ('Envoys were sent to Athens')

(Gildersleeve and Lodge 1968: 214). This appears to be problematical for those who propose a unitary function for each case (e.g. Jakobson 1936), and it encouraged various proposals concerning 'primary' and 'secondary' functions of cases (see e.g. de Groot 1939, Kurytowicz 1949). This added another dimension to the traditional characterization of 'case'.

Among 'local' cases many languages distinguish at least locative vs. allative vs. ablative. Yup'ik (Eskimoan) adds to this a vialis case:

(1) *-mi* locative ('location at, in, on')
-mun allative ('motion to, towards')
-mek ablative ('motion from')
-kun vialis ('motion through') (Mithun 1999: 138)

But there are also more extensive systems that distinguish variants of at least the first three of the types illustrated in (1) in terms of location relative to a reference point, involving dimensionality and orientation. Comrie (1981b: 210) associates six series of locative cases with the Kubachi dialect of Dargva (North-East Caucasian), three of which are given in (2):

(2)

locative	allative	ablative	
-ži-w/j/b	*-že*	*-ži-l*	'on'
-gi-w/j/b	*-gu*	*-gu-l*	'under'
-ta-w/j/b	*-ta*	*-ta-l*	'in front of'

(Each of the alternatives in the locative form is a marker in concord with an absolutive form.) These distinctions between 'syntactic' and 'local' inflections, and among the 'syntactic cases', are essentially a nineteenth-century development.[2]

Note too that the use of all of the 'syntactic' functions as definientia raises rather acutely the question of what is the independent content of such 'grammatical relations'. See further note 2.

[2] A further complicating, if not obfuscating, development, primarily again of the nineteenth century, is the proliferation of types of 'subject', associated with the observation that philosophers refer to 'logical subjects' that are marked by other 'cases' than the nominative. Padley notes, from the grammarian Alexander Hume (1612), 'an early use in grammar of *subiectum* for the logical subject of an utterance' (1976: 117). Hume doesn't apply the term to the argument of a verb but to *filius Socratis* ('son of-Socrates'), where we have 'a subject governed by a preceding adjunct'. By the nineteenth century distinctions are commonly being made among 'logical', 'psychological', and 'grammatical subjects'. This multiplies the problem of the nature of subjecthood. Does the 'grammatical subject' have any content apart from 'what is expressed by the nominative'? What do these different 'subjects' share apart from the label itself?

The displacement of the label 'subject' has also taken place, more recently, in another dimension than that of functional types. Consider e.g. Chomsky (1981: 209–10), with regard to 'the basic structure of S', as given in his (66):

(66) NP INFL VP, where INFL = [[±Tense], (AGR)]

where 'AGR = PRO is obligatory with [+Tense] and excluded with [–Tense]':

…let us introduce the term 'SUBJECT' having the following sense: the subject of an infinitive, an NP or a small clause…is a SUBJECT; AGR in (66) is a subject, but NP in (66) is not if INFL contains AGR.…Thus we take the SUBJECT to be the capitalized element in (67):

(67) (i) John [$_{\text{INFL}}$ past AGR] win
(ii) he wants (very much) [for JOHN to win]
(iii) he believes [JOHN to be intelligent]
(iv) [JOHN's reading the book] surprised me
(v) he considers [JOHN intelligent]

978–0–19–920647–6 09-Spencer-ch08 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 123 of 135 May 16, 2008 18:57

However, within this 'morphological' tradition, a viewpoint that was more familiar prior to the nineteenth century was to regard only the nominative (as well as the vocative) as exceptional in not marking a semantic relation. The origin of the term 'case' (from the Latin translation of the Greek πτωσις) has aroused much speculation, but, however formulated, the usual traditional division recounted by Meiklejohn (1892: 19, n.(i)) is persistent:

> The word **case** is from the Latin *casus*, and means a falling. The old grammarians regarded the nominative as the *upright case*, and all the others as *falling* from that. Hence the use of the words *decline* and *declension*. (Of course the nominative cannot be a real case, because it is *upright* and not *falling*.)

The name of each of the 'real cases' (in Latin, accusative, genitive, dative, locative, ablative) is suggestive of (what was seen as) its primary, semantic role (though these Latin terms sometimes mistranslate the Greek names): these cases are semantically grounded.[3]

Since the nominative lacks a semantic role such as the terminology attributes to the others, this raises a question as to what kind of role it plays in the structure of predications. Later, sporadically, but increasingly from the sixteenth century (Padley 1976: 52), the term 'subject' is adopted from logic among humanist grammarians and their successors: the nominative is the case of the 'subject'. (And, as we have seen, such non-semantic functions were later extended to some other inflections.) The Port-Royal grammarians (Lancelot andArnauld 1660: ch.VI) distinguish the nominative as not a 'real case', but as that from which the (other) cases are formed, but they also give its principal use as expressing the subject of the proposition. The consequences for grammar of the importation from logic of 'subject' have been the subject (indeed) of controversy ever since (recall note 2).

The important point is that the conception of non-nominative cases as differing from the nominative in expressing a semantic relation is maintained by many scholars, even in the nineteenth century, and in opposition to the views of Rumpel and many others since, although the latter indeed became prevalent in the twentieth century. The nineteenth-century localists (such as Wüllner 1831), in particular, regarded the nominative as exceptional, but analysed all the other cases as expressing spatially-based notions. Consider Wüllner's classification of the Latin case system, represented by Hjelmslev (1935: 39) as in Figure 8.1:

Here 'SUBJECT' fails to refer to the prototypical exemplar of the traditional 'subject'. And this disjunctive definition not merely changes the idea of how subjecthood is expressed but also makes its independent content even more mysterious – apart from what might be deduced from its alleged role in the formulation of often ephemeral theory-internal 'principles'.

[3] The 'nominative' terminology, however, harks back to Aristotle's name for this 'case': it was simply the citation form of the noun. Usually even more firmly excluded as a 'real case' is also the form of address, the vocative, despite its similarity in morphological expression to 'real cases'. For an analysis in localist terms, see Anderson (2007: §8.4).

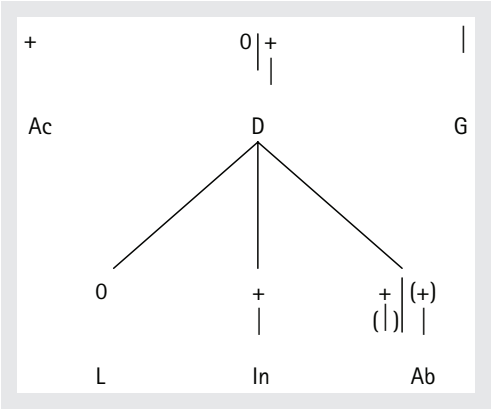


Figure 8.1.

The cases are Ac(cusative), D(ative), G(enitive), L(ocative), In(strumental), Ab(lative) – though some of these are at best only marginally discriminated in Latin. Accusative is positive with respect to the dimension of direction, Genitive is negative, while Dative is either neutral or complex (both goal and source). The others are specializations of this last. The fact that Wüllner's theory is here articulated in relation to a particular language, and whatever the merits of this particular articulation, should not be allowed to obscure its cross-linguistic potential.

I take up the idea of localism below; but there are other relevant earlier developments that are not reflected in the prevalent twentieth-century views on case, but are relevant to the development of localist case grammar. This finally brings us to the 'other hand' anticipated initially: that is, we turn now to the other main attitude to the study of case that had evolved by the modern period: as observed initially, the twentieth century also largely adopted, as well as an acceptance of the importance of grammatical relations, the dominant nineteenth-century usage that 'case' was a matter of morphology.

A different conception concerning 'case' and the relations they express had already emerged particularly among the humanists, however: there it is acknowledged that the relations expressed by case inflections may also be expressed in other ways. These relations are often expressed in Latin by a combination of preposition and inflection. Thus, the locative relation expressed by the ablative alone in (3) requires a preposition in other circumstances, as illustrated by (4):

(3) *Menippus . . . tōtā Asiā disertissimus*
Menippus . . . all:ABL Asia.Minor:ABL the.most.eloquent
'Menippus, . . . the most eloquent man in all Asia (Minor)'

(4) *Copias in castrīs continent*
riches in forts:ABL they.keep
'They keep riches in forts'

978–0–19–920647–6 09-Spencer-ch08 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 125 of 135 May 16, 2008 18:57

Moreover, whereas *in* plus the ablative in (4) expresses (non-directional) location, *in* plus the accusative is directional:

(5) *In Graeciam pervēnit*
in Greece:ACC s/he.arrived
'S/he arrived in Greece'

Differentiation of these locational relations requires the co-operation of preposition and inflection.

In other languages the relations expressed by case inflections may be signalled by adposition alone, and indeed by position alone, as well as by alternative morphological means. In recognizing this diversity of expression, the Port-Royal grammarians inherited this humanist tradition. Chomsky gives a succinct account of their views on 'case' (1966: 44–5):

> Chapter VI of the Port-Royal *Grammar* considers the expression of these relations in case systems, as in the classical languages, or by internal modification, as in the construct state in Hebrew, or by particles, as in the vernacular languages, or simply by fixed word order, as in the case of the subject-verb and verb-object relations in French. … Notice that what is assumed is the existence of a uniform set of relations into which words can enter, in any language, these corresponding to the exigencies of thought. The philosophical grammarians do not try to show that all languages literally have case systems, that they use inflectional devices to express these relations. On the contrary, they repeatedly stress that a case system is only one device for expressing these relations.

This idea of case as a **variant expression** among others of relations between words was widely adopted among universal and 'philosophical' grammarians. And such a view is enshrined in some early-twentieth-century discussions of the category of 'case', such as Hjelmslev's (1935, 1937), which takes an 'inclusive' view of 'case'. Hjelmslev recognizes this (1935: 24) as part of the tradition he inherits.[4]

The recognition of the 'positional' expression of 'case' relations is perhaps the final component in this multifaceted 'humanistic' view of the expression of 'case' relations that was adopted by Fillmore and others in the mid-60s, rather than the morphological tradition that elsewhere has come virtually to monopolize the study of 'case' since the nineteenth century. Fillmore (1965, 1966, 1968), however, maintains the distinction among (derived) 'subject', 'direct object', and 'indirect object', unlike the localists (including Anderson 1984); in this respect Fillmore aligns himself with the prevalent nineteenth-century tradition.

[4] But he mistakenly attributes the initial observation of the equivalence of preposition and inflection to Bernhardi (1805). As well as recognizing such a more inclusive view of case, Hjelmslev, by rejecting proposition (B) below, also (untypically for a localist) includes the nominative in his localistic analysis.

Let me try to sum up this brief sketch of what was in reality a complex non-linear history. I suggest the following characterization of 'case in traditional grammar', framed in a more recent terminology, but I think true to intention:

Case in traditional grammar

(A) There is a grammatical category whose content serves to differentiate among 'semantic roles', that is, the different 'roles' that may be assumed by participants and circumstantials in the scene described by predications containing instances of this category. These 'roles' have included 'place', 'logical subject', etc.; but for the localists they are all to be given a spatial interpretation.
(B) This category also differentiates a 'role' or 'roles' that is/are in some sense 'grammatical', such as '(grammatical) subject'; for the localists this is the only such 'role'.
(C) According to one tradition, this category may be expressed in various ways: by various morphological means, notably nominal (but also verbal) inflection, by adposition, by position.

(B) is rejected as a whole by Hjelmslev and (C) by the nineteenth-century tradition that has had most subsequent influence. Largely because of the latter rejection and adoption of the Rumpelian view of grammatical relations, and the acceptance of this conjunction in 'school' grammars, the particular combination of insights embodied in (A)–(C) was lost to twentieth-century linguistic theorizing, and in the main has remained so. This is not merely a terminological matter, but involves the unity or otherwise of what is expressed in these different ways, and its relative centrality to syntax.[5]

[5] The loss of the 'humanist' tradition is also reflected in the development of the idea of the 'autonomy of syntax' in place of traditional 'notional grammar' (see Anderson 2005 on the latter development, as well as the remarks in note 7). Such a view of syntax inhibits any unified implementation of the 'humanist' view; rather, the phenomena concerned are dispersed over various interface interactions.

Such a formulation of the 'humanist' tradition as I have given underlines the enormity of the misrepresentation of history that is embodied in Chomsky's equation, in the sentence that immediately precedes the quotation given above, of 'deep structure' with the concerns of the Port-Royal grammarians (1966: 44):

> The identity of deep structure underlying a variety of surface forms in different languages is frequently stressed, throughout this period, in connection with the problem of how the significant semantic connections among the elements of speech are expressed.

In Chomsky's own words, 'Chapter VI of the Port-Royal *Grammar* considers the expression of these relations in case systems', and the sentence I have just quoted apparently (correctly) identifies 'these relations' as 'semantic connexions among the elements of speech', 'corresponding to the exigencies of thought' (Chomsky 1966: 45), not autonomous syntactic 'deep structures'. The introduction of such a concept simply serves to obscure the significance of what the traditional study of 'case' had achieved. The illegitimate equation is made blatant by Chomsky's comment on work of Arnauld: 'he points out repeatedly that what is affirmative or negative "in appearance" may or may not be in meaning, that is, in deep structure' (1966: 44). This misrepresentation is pursued at some length: Chomsky attributes to these grammarians 'a belief that deep structures are fundamentally the same across languages' (1966: 44), for instance. Such a belief (by anyone) is now generally admitted to be meaningless, but the belief that these earlier grammarians were precursors of transformational grammar persists despite this.

It is this retreat from the traditional 'humanist' view embodied in (C) that Fillmore and others set out (consciously or not of that tradition) to reverse in the late 60s and early 70s of last century, at the same time as asserting the primacy of semantic over grammatical roles. The details of the non-localist case-grammar formulations of Fillmore and others need not concern us. Nor am I concerned with the wide range of objections to these formulations. Many of these objections seem to me, even in the context of the time, misguided, for reasons I have provided in a number of places (e.g. Anderson 1977: particularly ch. 1; Anderson 2006: particularly chs. 4, 5, 9). But one persistent set of problems relates to a particular failure on Fillmore's part to fully satisfy the implicit demands of the traditional (inclusive) concept of 'case' summarized above; and localist case grammars seek to modify this. The problems concern the identification of what determines the set of 'cases', i.e. the 'category of case', and of individual 'cases' within that set; and these deficiencies are indeed still to be found in most frameworks that appeal to semantic (or 'thematic', or 'case') relations (though they are not resolved by simply banishing these relations from status as primitives).[6]

These problems arise from the absence in Fillmore's proposals of a theory of 'case relation'. Notice that, according to Fillmore's notation, and as an iconic manifestation of this absence, the 'case relations' do not even belong to the same category; his 'Objective' and 'Agentive' are apparently atomic labels for individual 'case relations'. Hjelmslev formulates the problem of delimitation that derives from the lack of a theory of 'case' as follows (1935: 4):

Délimiter exactement une catégorie est impossible sans une idée precise sur les faits de signification. Il ne suffit pas d'avoir des idées sur les significations de chacune des formes entrant dans la catégorie. Il faut pouvoir indiquer la signification de la catégorie prise dans son ensemble.

[6] Some of these suggested problems, however, including some that are still widely cited, are specious. Dowty (1989), for instance, produces 'three recurrent problems' allegedly arising from the conflict between the 'argument-indexing' role of 'case relations' and their semantic characterization. One of these is the 'problem' presented by pairs like *buy* and *sell* (a 'problem' already resolved in e.g. Fillmore 1972). Concerning these pairs, Ackerman and Moore (2001: 24), citing Dowty, claim that 'on an intuitive level one would assume that, e.g., Max is an AGENT and Mary is a RECIPIENT in both the [(i)] and [(ii)] sentences':

(i) Max sold the piano to Mary for $1,000

(ii) Mary bought the piano from Max for $1,000

Appeals to 'an intuitive level' are always suspect; linguists' structural intuitions (which are of diverse origins, including their own particular theoretical training) are not evidence. Anderson (2006: §5.3), who also comments there on the other alleged 'problems', points out that, in the present case, it doesn't follow from the assumption that the 'vendor' in (i) is an Agentive, as source of the immediate action described by the verb, that the 'vendor' in (ii) is also presented as an Agentive. In the latter instance, it is rather the 'customer' that is presented as the source of the immediate action, even though the same 'real-world' event may be being referred to by (i) and (ii). To maintain otherwise is to succumb to 'the objectivist's misconception' (DeLancey 1991): linguistic constructs do not represent the 'real world' but our perception, interpretation, and representation of it and other worlds.

The lack of a theory of 'case' betokens lack of full commitment to pursuit of the traditional programme. Hjelmslev himself proposes a localist account of 'case'.

It is time, indeed, to take up localist case grammar, but we should firstly note another relevant problem that arises with Fillmore's suggestions, however. This follows from the fact that, though he makes a distinction between 'case' and its expression, between 'case relation' and 'case form', it remains unclear how exactly the diversity of expression is to be allowed for – except for the suggestion that marking of grammatical relations by position in English is associated with loss of semantic-relation etiquette.

The frameworks of Anderson (1971) and (1977) and Böhm (1982) adopt (variants of) a localist theory of case, as well as the 'humanist' view defended by Fillmore of (morphological) case as one variant in expression, one kind of 'case form'. As we have seen, localism involves the claim that spatial distinctions exhaust the domain of 'case', or rather the relations (equivalent to Fillmorean 'case relations') expressed by 'case' in the inclusive (variable-expression) sense. These I shall refer to as **semantic relations** (laying aside for the moment the status of 'subject').

Anderson (2006: §6.2) offers the following implementation of this idea, in terms of the exhaustive set of semantic relations given in (6). Consider, firstly, the set in (6a), consisting of all the contrastive members of the 'category of case':

(6) a. abs source loc loc{source}
⇓ ⇓
b. {goal} {goal}

The system is based on **semantically grounded** (simplex) features loc(ative) and s(ou)rc(e), and abs(olutive), where this last is defined as the absence of either (or both) of the others, so 'neutral'. They are semantically grounded in that both semantics and distribution are essential to their identification.

Example sentences are given in (7)–(10):

(7) Bill tripped.
abs

(8) Bill is at the window.
abs loc

(9) Betty betrayed Bill.
src abs
⇓ = abs{goal}
{goal}

(10) Bill slid from the door to the window.
abs loc{src} loc
⇓ = loc{goal}
{goal}

978–0–19–920647–6 09-Spencer-cho8 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 129 of 135 May 16, 2008 18:57

An absolutive occurs in each, alone in (7), with a locative in (8), as shown in the 'glosses'. In (9) a source occurs with absolutive; it is not 'loc{src}', where source is a **second-order** feature, but simply source, so it represents the source of the action, rather than a concrete source; action is a domain that is characterized by spatial relations. In the presence of this source, the *Bill* element in (9) is assigned the second-order feature 'goal', as shown in (6b) and (9). Similarly, in the presence of a locative source, such as *from the door* in (10), a bare locative is interpreted as a goal. This gives us a locative system such as is realized morphologically in Yup'ik (if we allow for vialis as a combination of loc{src} and loc{goal} associated with a single element): recall (1).

(10) is of course ambiguous between an agentive and a non-agentive reading. *Bill* therein is represented either as in the 'gloss' to (10), i.e. simple absolutive, or as a combination of source and absolutive:

(11) Bill slid from the door to the window
abs,src loc{src} loc

This thus deploys the Gruberian dispensation of associating multiple semantic relations with a single element. *Bill* in (11) fulfils a **compound role**, involving more than one first-order relation, each of which has semantic and syntactic consequences.

Multiple relationality underlies other localist interpretations of apparently non-spatial relations, such as the 'experiencer' in (12):

(12) Bill knows that song
src,loc abs

Bill is interpreted as the location of the scene expressed by the predicator, as shown in its semantics and syntax (e.g. Anderson 2006: §5.4.3). But *Bill* is also the source of the scene, here an experience rather than an action (which latter is characterized by the presence of a source that is not combined with locative), and this is again reflected in the syntax. Notice that neither feature associated with the role of *Bill* is second-order, they have equal status; *Bill* has a compound role. As with the concrete locationals in (8) and (10), (12) has directional equivalents, illustrated by (13):

(13) Bill learned that song from Betty
src,loc{goal} abs loc{src}

Here *Bill* is the goal of an experiential transfer.

I should emphasize that for a localist it is not a question of the plausibility of individual localist analyses of some 'abstract' relations; but the localist hypothesis exhausts the domain of 'case'. All semantic relations are to be interpreted localistically. Moreover, the only grammatical relation (in a subject-forming language) is the neutralization of semantic relations we call 'subject'. Otherwise the distribution of case forms and their cross-linguistic identification (as 'accusative', or 'dative', etc.) is determined by semantic distinctions (recall Figure 8.1, and see for further, more detailed illustration Anderson 2006: §6.1). '(Direct) Object' and 'Indirect Object'

are incoherent notions. The pervasive appeal to them by syntacticians is largely a product of an educational system whose grammatical ideas (in this regard) are based uncritically on the nineteenth-century Rumpelian tradition.

I have interpreted (13) as non-agentive. But again it can be given an agentive interpretation, as is normally the sentence in (14):

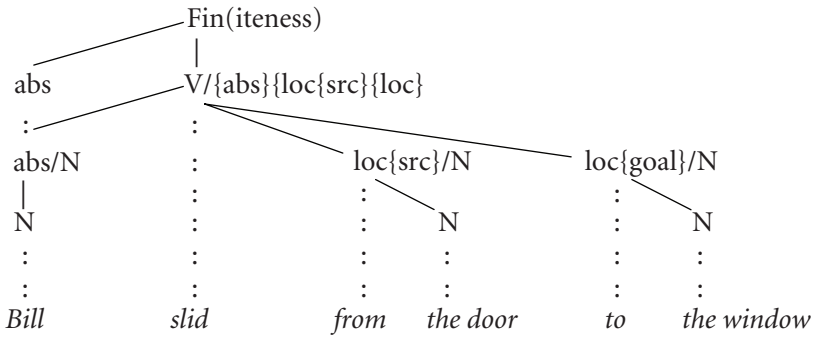
(14) Betty taught Bill that song

But these agentive sentences involve further considerations, in imposing an agentive structure on a directional. This is made possible by a complex lexical verbal structure. In order to investigate this, however, we must take up the representation of both syntactic and lexical structure as relations of **dependency**. At this point what we have illustrated is a proposal that the content of the category of 'case' is exhausted by the set of relations in (6).

In localist case grammar the structure of sentences is expressed by dependency trees; that is, the tree representations represent in graphic form **head-dependent** relations, rather than constituency relations, which are regarded as derivative of dependency. Heads are associated with **complements**, the elements whose presence in the more inclusive structure is licensed by the head, and with **modifiers**, which optionally extend the structure projected by the head; the head governs these other elements.

Thus, *to* in (10) licenses the following nominal, and is in turn licensed by the verb, as shown in the structure governed by *slid* in (15):

(15)



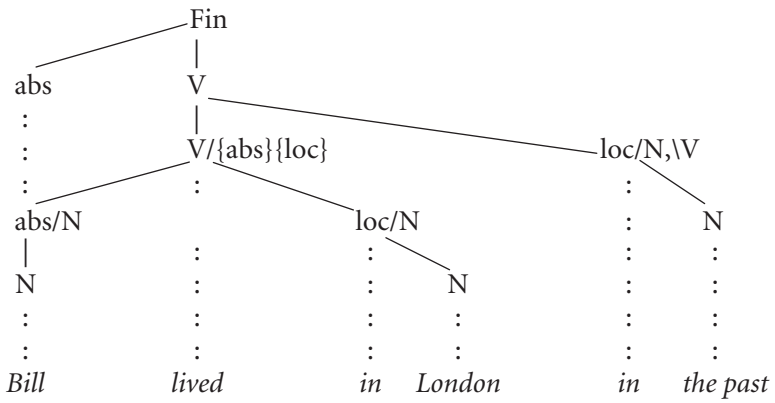
Leave the rest of the structure in (15) out of account for the moment. Likewise, ignore the placement of *Bill* to the left of the verb; this is associated with the hierarchy of semantic relations that determines subject-selection, whose role we shall return to.

In each case the set of complements, or valency, of the head is shown to the right of the slash, '/'; that for the verb is complex; and that for the semantic relations is the default, and can be given by redundancy. The head appears at the upper end of a solid line, representing the dependency arc linking head and dependent;

and the categories that label the nodes in the tree are associated with a lexical item (though I have not spelled out the internal structure of the N construction – see e.g. Anderson 2007); these associations are indicated by the discontinuous lines.

A modifier, on the other hand, seeks a head to modify, as represented in (16):

(16)

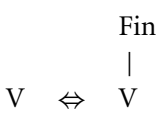


Here, the modifier *in the past*, as well as being marked for a complement ('N'), also signals that it seeks a head of category 'V'; this is the import of the backslash, '\'. It introduces above such a V a replica of that category, which gives the extra layer in (16); the extended structure remains verbal.

Sentence structure is built up in general on the basis of these specifications of complementation and modification, which are given in the lexicon. The internal structure of the complex verbs in (13) and (14) is also given in the lexicon. It is thus time to confront lexical structures in general. Consideration of the undiscussed parts of (15) and (16) introduces some relevant considerations, as well as clarifying our notion of semantic relation.

The dependency relations between the categories associated with different lexical items involve **adjunction**: the head and dependent are lexically disjoint. There are vertical solid lines in (15) and (16) that also represent a dependency relation, but in this case internal to a single word. Most of these internal dependencies are given in the lexicon, in the entry for the item concerned or by redundancy. So, any V can be expanded as in (17):

(17)



The upper V in (16), on the other hand, is introduced in the syntax to satisfy the modifier. These are the two ways in which **subjunctive** (as opposed to adjunctive)

dependency can be introduced. Words can thus be categorially complex, and each component category makes and satisfies syntactic demands.

Complex categoriality also allows us to accommodate systems of adpositions and case inflections that exceed what can be allowed for by the distinctions offered by (6). As well as the locational and directional distinctions allowed by (6) we also find, subordinate to them, distinctions of dimensionality (e.g. *in* vs. *on*) and orientation (*above vs. below*), marked morphologically in the examples in (2). As is suggested by the history of some of these, and by the evident complexity of their equivalents, adpositional (*into*) or morphological (as in (2)) or indeed 'periphrastic' (*in front of*), in many languages, they are categorially complex, in involving a 'dimensional' nominal subjoined to a simple semantic relation, as in (18):

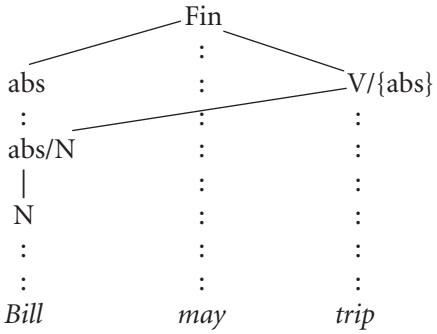
(18) loc
|
N

We have complex 'cases', which govern 'ordinary' nouns. Anderson (2006: §8.4) discusses such distinctions in relation to Hjelmslev's analysis of complex morphological system of Tabassaran.

The absolutive nodes in (15) and (16) are also associated lexically with a subjoined N; all Ns can be so expanded in parallel fashion to (17). This allows for semantic relations to be expressed by an independent word, as with the locatives in (15) and (16), or by position, as with the absolutives there, or morphologically, as with some of the arguments in the Latin examples (3)–(5). The subjunction/adjunction distinction is thus the crucial component in allowing for the variety of types of 'case form': the semantic or 'case' relation may thereby be expressed as a separate word or as part of a word, and thus inflectionally or by the position of that word. Such words are categorially complex.

This kind of distribution for a category is associated with **functional** status. Compare the finiteness category, which can have a lexically subjoined V, as in (15) and (16), or an adjoined predicator, as in (19):

(19)



Finiteness is a functional category. The semantic relations also belong to a functional category that Anderson (1997, 2006) calls **functor**; this is the name for the category of 'case'. The particular relations themselves are secondary features of this category, so that we should represent *from* as F{loc{src}}, with a primary (categorial) feature, a secondary, and a higher-order secondary.

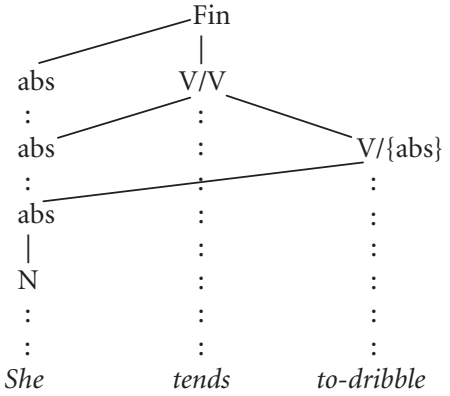
Finiteness, like other predicators (I have not characterized this status for it here), can take dependent functors. However, the attentive reader will have observed that finiteness is not subcategorized for these dependent absolutives. They are introduced at the lexicon–syntax interface by a redundancy that provides any predicator lacking an absolutive with such a dependent. This redundancy satisfies, whatever else, the logical requirement that a predicator must have an argument; indeed, it can be regarded as a grammaticalization of this requirement. Such absolutives, termed **free absolutives** by Anderson (2006), which do not introduce a semantic argument, may be satisfied as to their own complementation (functors govern Ns, as a default) by a 'dummy' N, as in (20):

(20) It rained
We hot-footed it to the square
It was expected that we would be late

In other circumstances, their valency is satisfied by **argument-sharing** with the semantic relation that is highest on the **subject-selection hierarchy** in a dependent (non-finite) V, as with the finiteness absolutives in (15) and (16), and as represented in (22), for (21), where I ignore the status of *to*:

(21) She tends to dribble

(22)

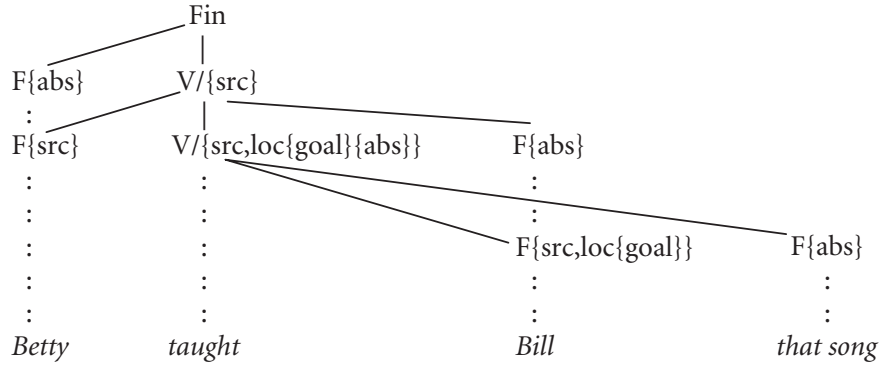


She is shared among the subcategorized-for absolutive of *dribble* (which is by default highest on the subject-selection hierarchy in its predication), the free absolutive of *tends*, which is a V that takes only V as an argument, and the free absolutive of finiteness.

Argument-sharing is only available to free absolutives, and it avoids appeal to 'raising' and other movements (or their descendants). While dependants in English are generally serialized after their head, the free absolutive of finiteness comes before. This formulation, involving argument-sharing between the free absolutive of finiteness and a selected functor in the dependent predication and placement to the left of finiteness, defines the core of syntactic **subject formation**. And it means, because of argument-sharing between the free absolutive and different subcategorized-for roles, that the position is associated with neutralization, a neutralization conferred by this association with the higher free absolutive.

Much of this can be illustrated by the representation for (14), which has purposely been delayed until now:

(23)



The verb is categorially complex not merely in being subjoined to finiteness but also in involving a directional predicator subjoined to a causative. The finiteness free absolutive shares the *Betty* argument of the causative source; the free absolutive of the causative shares the experiencer of the experiential directional. The source and the experiencer ({src,loc}) are hierarchically highest (outranking absolutive) on the subject-selection hierarchy in their respective predications. (See further Anderson 2006, particularly ch. 13).

Such work on localist case grammar in the last decades has suggested that it was an error to try to express a grammar of case in transformational terms. As well as introducing undesirable power and complexity, syntactic derivationality is superfluous in a framework that has available to it the expressibility made available by an implementation, along with dependency structures that permit restricted argument-sharing, of the traditional insights concerning 'case', such that semantic relations are basic to both the lexicon and the syntax, and that morphology is only one expression-variant for 'case' or semantic relations. The work alluded to here thus involves developments in ideas about syntax that move in the opposite direction from the kind of 'abstractness' to be found in the tradition associated with Hale and Keyser (2002). This is enabled by the assertion of the primacy of semantic

relations over grammatical, not merely by the less interesting claim involved in simply recognizing the relevance of both, and that they are in some way 'linked' or 'interactive'. This area again is discussed in some detail in the later chapters of Anderson (2006) – and see too Anderson (1992, 1997), and Böhm (1998, 1999) and other earlier work referred to there.[7]

[7] These chapters of Anderson (2006) are also concerned with related work that has broadened the significance of localist case grammar in an even more drastic way. For it is not only that localism applies to other categories than 'case' (such as aspect and quantification). Localism itself is but one aspect of the implementation of a more general assumption concerning the nature of grammatical categories in general, and indeed concerning the basis of syntax. A case grammar is (part of) a notional grammar, a grammar (again traditional) whose generalizations are grounded in semantics and/or pragmatics, as phonology is grounded in phonetics. Of course, syntax is not semantics/pragmatics, and phonology is not phonetics; but they are partial conventionalizations of these. Grammatical regularities certainly may become what Anderson (2006), for instance, refers to as 'routinizations', regularities that show decreased sensitivity to meaning (as with routine 'second-position' for verbs). But these are best understood and represented within components of the grammar that are grounded in substance. On such a view, both meaning and distribution are necessary for the identification of syntactic categories in general, not merely that of 'cases'.

CHAPTER 9

CASE IN COGNITIVE GRAMMAR

SILVIA LURAGHI

9.1 Introduction

The present chapter is devoted to cases in Cognitive Grammar (henceforth CG). Cognitive approaches to cases are numerous, and by no means restricted to CG. Some other cognitive approaches are treated in separate chapters in this handbook (e.g. Wierzbicka, Chapter 10). In this chapter, I will start by framing views on grammatical forms common to various cognitive approaches in the framework of earlier research on the same topic (§ 9.2); I will then briefly survey the context out of which CG originated (§ 9.3); and proceed to illustrate more specifically relevant assumptions within the CG framework (§ 9.4). Finally, I will provide some diachronic evidence for the CG approach to cases (§ 9.5).

9.2 Traditional approaches

In cognitive approaches, language is viewed as an integral part in the whole of human cognitive capacities, rather than a separate module; furthermore, it is

assumed that language can be externally motivated. A consequence of this assumption is that meaning is viewed as being pervasive. According to cognitive approaches, grammatical forms are meaningful elements. This assumption is not new to linguistics: historical linguists of the nineteenth century have devoted long discussions to topics such as the meaning of the accusative, etc. In early historical linguistics, that cases had meaning was an assumption that hardly needed to be argued for. Meaning was also viewed as the major explanation of change: for example, case syncretism (which, in a diachronic perspective, indicated merging of different cases over a period of time, and not synchronic polysemy)[1] was viewed as dependent on similarity in the meaning of the cases involved, in very much the same way as in CG (see § 9.4.5.1 and 9.5.1).

Works devoted to cases flourished in the late nineteenth to early twentieth century, and case meaning was among the favourite topics of the Neogrammarians. These works were in large part theoretically naive. scholars did not usually adduce general principles that could support their assumptions, partly because, as already remarked, the fact that meaning had a central role in grammar was considered self evident, as was the idea that language could be externally motivated. As an example, one can quote Delbrück (1867: 50–1), who explains the extension from comitative to instrumental meaning as due to the fact that 'the instrument is the entity in whose conjunction we bring about an action', thus foreshadowing Lakoff and Johnson's 'companion metaphor' (see Lakoff and Johnson 1980: 135), or Wackernagel's explanation of the frequent polysemy involving location and instrument, which he connected with the existence of specific instruments which are also containers, like e.g. quivers (see Wackernagel 1922: 304–5).[2] Because the work of the Neogrammarians was grounded on deep knowledge of the languages they investigated, it would be worth trying to benefit from their insight as much as possible.

In the 1930s, the tradition of research on case meaning was taken over by the European Structuralists, who produced a significant number of studies on this matter, this time with an elaborated theoretical background. Among them the most influential remain Hjelmslev's *La catégorie des cas* (1935), and Jakobson's *Beitrag zur allgemeinen Kasuslehre* (1936).

Jakobson (1936) still constitutes the basis for research on Russian cases. Jakobson did not see himself as outside the lines of traditional research, but rather framed his work in earlier ones. The main difference between his (as well as Hjelmslev's) work and the Neogrammarians' lies in Jakobson's assumption that cases built a system, and that consequently their meanings had to be investigated

[1] See Luraghi (2000) on the meaning of 'syncretism' in the tradition of historical linguistics.

[2] Delbrück expressed his views on the relation between comitative and instrumental in reference to the Old Indic instrumental case, which can encode both semantic roles. Wackernagel's discussion of location and instrument does not concern a case marker, but a derivational suffix, Ancient Greek *-tro-*, which can form both instrument nouns and nouns denoting location.

978–0–19–920647–6 10-Spencer-ch09 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 138 of 150 May 16, 2008 18:57

in relation to each other. A major point in Jakobson's study of cases (as well as in many other contemporary or earlier studies) is constituted by his interest in finding a 'general meaning' of each case (*Gesamtbedeutung*). Jakobson quotes Hjelmslev, who thought that the 'basic meaning' (*signification fondamentale*) was a concept that explained all concrete uses of a certain case marker. This approach proves somewhat rigid when tested on a large mass of data, but it remained the most effective, until the model of radial categories was worked out by Lakoff (1987).

The Structuralists' approach was certainly less naive than traditional approaches from the theoretical point of view, but, seen from a CG perspective, it had the disadvantage of focusing the analysis of case meaning on oppositions within the system, rather than on the substance of the meaning itself. Note however that this does not per se mean that grammatical meaning is conceived as substantially different from lexical meaning, since a goal of structuralist semantics was to show that lexical meaning could also be organized in the same way.

9.3 Early cognitive approaches: 'deep' cases

In early generative grammar, little if no attention was paid to morphology. The phonological component was conceived as linking directly with the syntactic component, so that bound morphemes such as case markers were seen as mere by-products of mapping rules that produced the surface structure of sentences. The idea that especially inflectional morphology was not worth being investigated was so strongly entrenched in American theoretical linguistics that even after a number of cognitively oriented linguists started reacting against generative grammar it took some time for them to understand its real importance.

The prehistory of CG dates back to the 1960s with generative semantics. In 1968, the word 'case' was brought to the forestage by Fillmore's influential paper *The case for case*. What Fillmore called 'cases' or, with a more accurate label, 'deep cases', corresponded to what is nowadays known as semantic roles, and had nothing to do with case markers. Even if cognitive semantics highlighted the importance of meaning, the fact that all grammatical forms had a meaning that could explain their usage took one more step to become commonly accepted.

Interestingly, research on the semantics of cases (intended as 'surface' cases) continued directly out of Jakobson's work especially among scholars working in the Slavic linguistic tradition, even when they were aware of Fillmore's theory of 'deep cases', as shown for example in Wierzbicka (1980).

9.4 Cases in Cognitive Grammar

In the last decade, quite a lot of research has been devoted to cases within the framework of CG. In the following sections, I will briefly summarize some assumptions that must be kept in mind in order to understand the CG analysis of case meaning (9.4.1–4); I will then proceed to illustrate a number of recent studies devoted to cases within the CG framework.

9.4.1 Radial categories

Radial categories have been described in Lakoff (1987). They are a type of prototypical categories. Radial categories have a central subcategory, that displays all features relevant to the category. Non-central subcategories 'are not specialized instances of the central subcategory, but rather are variants of it . . . These variants are not generated by the central model by general rules; instead, they are extended by convention and must be learned one by one' (1987: 91). This means that one does not automatically know what the actual members of a radial category will be; the features of the central subcategory, however, albeit not determining which subcategories will be generated, explain why those that actually exist could be generated: 'The central model determines the possibilities for the extensions, together with the possible relations between the central model and the extension models' (ibid.).

The relevance of radial categories to the study of case meaning has been highlighted by Janda (1993), who especially pointed out the fact that radial categories are internally structured, thus offering a very insightful model for semantic extension. Especially in the case of highly polysemic items, such as cases, the fact that the various meanings can be related with a central model, but differently located within the complex structure of the category, enables one to 'seek as much detail in [one's] description as desired without endangering the integrity of the category. The network may become increasingly intricate, but by virtue of the fact that its structure is based on interrelationships, constant reference is made to the prototype and those members closest to it' (Janda 1993: 6).

By considering case meaning as having the structure of a radial category, one can avoid long lists of unrelated 'functions of cases', as in many reference grammars, while in the meantime not being forced to consider all meanings on the same plane and equally related to each other in such a way as to build a unitary fundamental meaning.

9.4.2 Polysemy vs. homonymy

A frequent trend in the structuralist tradition has been to set up various homophones in cases where grammatical forms are used in different functions.

A consequence of positing radial categories is that CG can minimize the number of homophones, and assume that polyfunctionality of a form is better explained as due to polysemy. In her analysis of the meaning of the genitive, Nikiforidou (1991: 196) further remarks that assuming structured polysemy, rather than homonymy, allows one to 'provide an explanation for the different status of each meaning (central vs. peripheral)'.

Under the assumption that polysemy is based on a structured network of relations among meanings, Langacker (1991: 379) can easily prove pointless a common argument against the idea that cases are meaningful, i.e. that it is impossible 'to isolate any single meaning that is clearly appropriate for a particular case in all its occurrences'.

9.4.3 The meaning of grammatical forms

It is commonly said that, whereas lexical forms convey lexical meaning, grammatical forms like bound morphemes convey grammatical meaning. This may be a practical way to put it, but it is important to understand that two types of meaning referred to as grammatical and lexical are not qualitatively different: grammatical forms are meaningful in very much the same sense in which lexical forms are meaningful, the only difference being that the meaning of grammatical forms is more abstract.

Langacker (1991) has a number of illuminating remarks regarding case meaning (see also above, § 9.4.2). Focusing on cases other than the nominative and the accusative, Langacker remarks that 'It is quite evident that the oblique cases [like dative, instrumental, locative, ablative, etc.] make a definite semantic contribution in many (if not all) of their uses . . . If the subject- and object- marking cases depart somewhat from this model, it is only by virtue of the abstract nature of their meaning' (1991: 379–80).

This passage makes clear that what can look like a difference between lexical and grammatical meaning is only a matter of increasing abstractness: indeed this follows naturally from our knowledge about grammaticalization processes. According to grammaticalization theory, lexical forms may acquire new meanings through metaphoric extension, whereby the extension often implies that one specific feature of the original meaning is magnified, with a gestalt effect (see Lakoff 1977, Heine et al. 1991: 43), and other features become irrelevant. This loss of parts of the original meaning results in semantic bleaching: the meaning of a form becomes increasingly general and abstract, as for example in the case of Late Latin *casa* 'home', which has become the French preposition *chez* 'at', preserving only the feature of location out of the original meaning. Consequently, new grammatical forms rise out of lexical forms, and following semantic bleaching phonological autonomy may be lost, so that once autonomous forms may become clitics and later bound morphemes. In

the case of case markers, there is clearly a continuum between independent adverbs (often derived from nouns), adpositions, and cases, as shown by cross-linguistic evidence: different languages may have adpositions that encode functions encoded by case markers in other languages (for example, French has the preposition *de* where Russian has the genitive case).

Another argument against meaningfulness of cases, which is adduced by functional linguists of various orientation, is that 'case is often *governed* by a verb, preposition, or construction' (Langacker 1991: 379), and it is frequently thought that if an item is obligatory in a certain syntactic environment it does not convey meaning. As Langacker points out, however, obligatory does not mean meaningless: 'A morpheme's failure to provide independent semantic content does not imply that it is semantically empty, but only that its contribution is redundant; all composition involves semantic overlap, and full overlap is an expected limiting case' (ibid.).[3] The importance of this remark cannot be overestimated: that governed forms have no meaning has been a commonplace argument in the description of cases in recent decades, for example in the framework of valency grammar and of Simon Dik's Functional Grammar (see for example Pinkster 1990). Because traditionally oriented research on the ancient Indo-European languages are based on the assumption that grammatical forms are meaningful, and that there is basically no difference between lexical and grammatical meaning (see § 9.2), the Functional Grammar approach to cases was held by its proponents as an advancement over the antiquated idea that cases had meaning (which indeed reached out to the origins of comparative linguistics, much in the same way as the theory of grammaticalization, see Heine et al. 1991: 5–8).

9.4.4 Trajector–landmark asymmetry

In CG, the relation set up by a case is regarded as an instantiation of the trajector–landmark (or figure–ground) asymmetry: an entity with a relational profile (the trajector/figure) is foregrounded with respect to another entity, which serves as reference point, and is backgrounded (the landmark/ground). The asymmetry is connected with salience of the two entities involved. In Langacker's words, landmarks 'are naturally viewed (in prototypical instances) as providing points of reference for locating the trajector' (Langacker 1987: 217), as for example in *the book on the shelf*: a shelf is a bigger and much more stable entity than a book, which can easily be moved. So it is so to say natural that the book is chosen as landmark

[3] Heine et al. (1991: 28) points out that 'whereas "concrete concepts" are autosemantic, . . . grammatical concepts have been described as . . . acquiring semantics by combination with other concepts': Langacker's remark on the role of semantic overlap can also apply to this distinction and show that while it may be true that the role of the context for the understanding of the meaning of a specific form is maximized with grammatical forms, it is also important with lexical forms, and it is essentially the same type of phenomenon, with only a difference in degree.

Au: Please provide figure legends for all figures in this chapter.

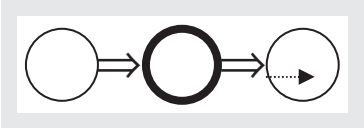


Figure 9.1.

(or ground). In an NP as Latin *domus patris* 'father's home', the genitive case of *patris* establishes a relation between a trajector (*pater*) and a landmark (*domus*), in very much the same way as English's. The trajector–landmark asymmetry is central to any type of relational predication in the framework of CG (see Langacker 1987: 217f.).

The specific case marker (or adposition, in languages without cases) profiles a certain relation between the trajector and the landmark. This can be represented as in Figure 9.1, from Langacker (1991: 404), which represents the meaning of instrumental *with* in an English sentence like *Floyd broke the glass with a hammer*. In this Figure, 'the schematic conception of an action chain . . . designates a participant characterized in terms of the *instrument* role archetype' (Langacker 1991: 405).

9.4.5 Some examples

In the introduction of her (1993) book on the Czech dative and the Russian instrumental, Laura Janda quotes a remark by Anna Wierzbicka, who in 1980 wrote that 'cases have fallen on hard times'. Looking back to the last twenty-five years, one can remark that things have changed dramatically. Research on cases has grown, in various theoretical approaches, mostly due to the progress of language typology, and its influence on linguistic theory. The CG framework proves particularly insightful for the study of cases, because of its theory of meaning, as expounded in the preceding sections. Consequently, there are now numerous CG-oriented studies on cases in different languages. In the following sections I will only survey a few, in order to point out what I think are the more relevant contributions CG can give to a general understanding of this matter.

9.4.5.1 *The genitive*

Nikiforidou (1991) devoted a long paper to the meaning of the Indo-European genitive, focusing on different languages (Classical Greek, Latin, Old English, and Medieval French). It may be argued that the sample could easily have been made more significant by the addition of some other languages in which the genitive plays an important role (Old English and Medieval French are certainly not among the best representatives of languages with cases in the Indo-European family); fortunately, however, Classical Greek and Latin offer quite typical examples in this

978–0–19–920647–6 10-Spencer-ch09 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 143 of 150 May 16, 2008 18:57

respect. Nikiforidou is not aware of traditional descriptions of cases that could have proven in accordance with her analysis (her critique of earlier studies basically refers to the mid-sixties Structuralists),[4] but apart from minor flaws, her description of the meaning of the genitive is a good example of how case meaning is structured.

Nikiforidou shows how different functions of the genitive are based on metaphorical meaning extensions. Starting with alienable possession, Nikiforidou argues that the meaning of the genitive can extend to inalienable possession, based on the metaphor 'parts are possessions', by which the possessor–possessee relation is mapped onto the whole–parts relation (1991: 170). A further metaphor, according to which 'wholes are origins' explains why in certain languages (notably Classical Greek) the genitive may extend to relations that are typically encoded by the ablative in other languages (1991: 173). Similar metaphors are shown to explain all possible meanings of the genitive, in accordance with the structured polysemy model (see § 9.4.2).

A merit of Nikiforidou's paper is that it also addresses the issue of diachrony. Cross-linguistic synchronic affinity between the genitive and the ablative shows that diachronic syncretism of the genitive and the ablative in pre-literary Greek follows 'a natural direction of change' (1991: 195), and in general 'the same semantic network set up for synchronic polysemy, can . . . be shown to be (at least partly) relevant to semantic change as well' (ibid.).[5] This is of course a tenet of traditional descriptions of case syncretism, see below, § 9.5.1.

9.4.5.2 *The Czech dative and the Russian instrumental*

In the Russian tradition of linguistic analysis, cases have never really been ignored, and among Russian cases, the instrumental has been a favourite of several scholars within various theoretical frameworks. Thus, Janda's (1993) book *A Geography of Case Semantics*, devoted to the Czech dative and to the Russian instrumental, draws on a long series of earlier analyses. In her description of case meaning, Janda makes use of the radial category model, showing that the complex and apparently random array of functions of the Russian instrumental can be captured by four image-schemas, that can be shown to build a network. The network has the 'conduit instrumental' (instrumental of means) as its centre. This is the prototype, and it is related to the 'instrumental of setting' (instrumental of space and time) and to the 'attributive instrumental' by either foregrounding or backgrounding the instrumental, or by focusing either on the participant or on the event. The fourth image-schema is the 'comitative and proximate instrumental'; it is derived in a

[4] There is a short reference to Jakobson (1936) (quoted in the 1984 reprint), in whose regard Nikiforidou (1991: 156) writes that she cannot 'evaluate the Russian examples', so her critique remains quite vague.

[5] See Luraghi (1987) for a similar explanation and further references.

similar way as the other two image-schemas, but it is related with image-schemas 2 and 3, rather than with the prototype (Janda 1993: 139–41).

To describe case semantics, Janda makes extensive use of the type of graphical representation used by Langacker and illustrated by means of Figure 9.1 above. Some examples are the schema of the indirect object dative and the free dative in Czech, given here in Figure 9.2(a) and (b) respectively (from Janda 1993: 53).

Janda's sophisticated analysis constitutes a very convincing example of how our knowledge about human cognition can implement the description of grammatical meaning, achieving new insight without necessarily rejecting all previous scholarship. Much to the contrary, Janda shows that her analysis of the Russian instrumental is in accordance with Jakobson's, even if the theoretical framework is different, and similar explanations are reached under quite different assumptions.

9.4.5.3 *German cases*

German has a rather reduced case system, mostly used for encoding grammatical relations; the meaning of German cases is more abstract than the meaning of Russian or Ancient Greek cases, and one would expect their semantic contribution

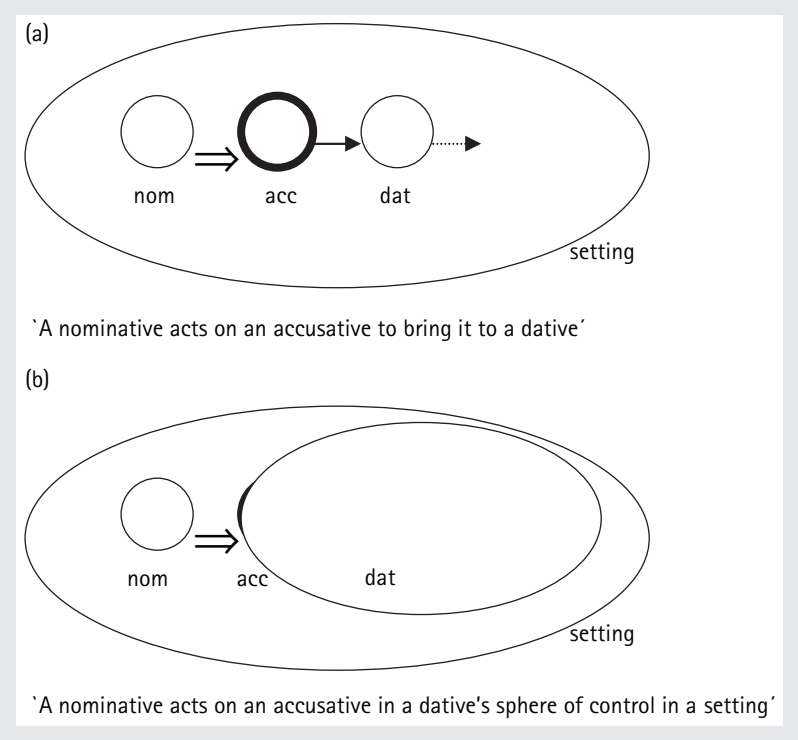


Figure 9.2.

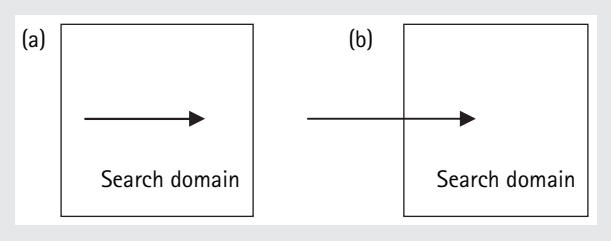


Figure 9.3.

to be highly conventionalized. Smith (1993) analysed the meaning of the accusative and the dative, both when used on the clause level and when used as complements of prepositions. According to Smith, the accusative case 'signifies, in its prototypical sense, the physical movement of a TR[ajector] along a path which makes contact with a significant aspect of the L[and]M[ark]' (1993: 533), while the dative variously signals a 'departure from ACC[usative]' (*ibid.* 534). Smith shows how all usages of the two cases can be brought back to these prototypes.

Smith's analysis of what he calls '2-way' prepositions is especially illuminating. Two-way prepositions are prepositions such as *in*, which may take either case, as in *Ich fahre in die Stadt* 'I drive into town (direction)' vs. *Ich fahre in der Stadt* 'I drive inside town (limits of motion)'. To explain the difference, Smith makes use of the concept of search domain: while the dative indicates that the trajector is wholly located in the preposition's search domain, the accusative indicates that the trajector moves along a trajectory which starts outside the search domain, and penetrates its boundaries. Ultimately, 'DAT[ive] and ACC[usative] signify unchanging state vs. change of state', two meanings that can be 'motivated as extensions from the prototypical senses of the cases' (1993: 540).

In order to represent the difference between the dative and the accusative with two-way prepositions in terms of search domain, Smith uses the image schemas in Figure 9.3 (from Langacker 1991: 403). Figure 9.3(a) represents the dative: the search domain is completely inside the landmark; while figure 9.3(b) represents the accusative: the trajector moves from outside the landmark to its inside.

9.4.5.4 *The Classical Greek accusative*

In Luraghi (2003) I analysed the meaning of the Ancient Greek cases and prepositions. In § 2.2.1 I argued that the accusative case basically signals total affectedness. This feature can explain both the function of the accusative as direct object marker, as well as possible alternation with other cases with certain verbs (for example, the alternation with the partitive genitive, which signals partial affectedness, or with the dative, which indicates that there is no change of state). Total affectedness also explains spatial uses of the accusative, as the accusative of direction, and, through

978–0–19–920647–6 10-Spencer-ch09 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 146 of 150 May 16, 2008 18:57

the notion of fictive motion (see Talmy 2000: 136), the so-called 'accusative of extension', which is used in reference to a stretch of space (not necessarily with verbs of motion) or of time. For example, the accusative occurs with verbs such as *apékhesthai*, 'to be far from', and indicates a distance.

In Classical Greek the accusative also had a wide adverbial usage, in which it encoded area, as in *pódas ōkùs Akhilleús* 'feet-ACC swift-NOM Achilles-NOM = Achilles swift-foot', or in expressions such as *diaphérein phúsin* 'to be different nature-ACC = to be different in nature'. This usage of the accusative is related to the accusative of extension, moved to an abstract plane.

As already remarked, the notion of total affectedness explains possible alternation between the accusative and the genitive, in occurrences such as *pínein oínon / oínou* 'to drink wine-ACC / wine-GEN = to drink (all the) wine / some of the wine'. On the plane of spatial relations, the same opposition motivates the occurrence of either case with prepositions: while the accusative conceptualizes a stretch of space without further implications on its internal structure, the genitive conceptualizes space as constituted by sub-parts. Accusative landmarks are continuous, while genitive ones are discontinuous, in the terms of Talmy (2000). Consequently, a trajector moving inside an accusative landmark does not follow a straight path, which can be described in a precise manner, but moves around randomly; much to the contrary, a trajector moving inside a genitive landmark moves along a clearly defined path, and can always be tracked down. This difference lies at the basis of alternations such as *dià dôma / dià dṓmatos baínein* 'through hall-ACC/GEN walk = to walk around in the hall / straight through the hall' (see Luraghi 2003: 169–71).

9.5 Diachrony

Perhaps the best argument in favour of a CG-oriented analysis of case meaning comes from diachrony. Even if historical consideration of semantic extension is not necessarily among the premises of a cognitive study of case semantics, when pursued it proves extremely important and fruitful, for several (related) reasons:

a) it can provide evidence for some assumed paths of semantic extension (see above, § 9.4.5.1);
b) it can help avoid too strong generalizations (e.g. that semantic change always goes in the same direction);
c) it can explain some otherwise hardly understandable phenomena, such as relative markedness.

978–0–19–920647–6 10-Spencer-ch09 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 147 of 150 May 16, 2008 18:57

9.5.1 Merging of cases

The Indo-European languages attest to various extents a process of reduction of their case system, traditionally called 'syncretism' (see fn. 1). It is generally assumed that cases did not simply disappear, but rather merged with other cases: this process can partly be demonstrated on morphological grounds, whereby the exponent of a certain case in a given language corresponds to the exponents of two or more cases in some other language(s).

As has also been pointed out by Nikiforidou (1991), quoted above, § 9.4.5.1, case syncretism can be shown to be in accordance with predictions about semantic extension that follow the model of structured polysemy purported by CG, for example in the case of syncretism of the ablative and the genitive. Radial categories can help us understand even more complicated instances of syncretism. As an example, I will illustrate the merger of the Indo-European dative, locative, and instrumental in Ancient Greek.

In the oldest Greek documents, the Mycenaean tablets, there is most likely only one series of exponents that continue the Proto-Indo-European locative and dative. The two cases had merged, as they did in some other languages (e.g. Hittite). In the languages that preserve it, the locative indicates a location where an event takes place. As opposed to other local cases, the locative denotes a static relation. The dative prototypically denotes experiencer, i.e. the semantic role of a human being who is affected by an event, but does not undergo a change of state. Again, the relation is static: this is the common feature of the locative and the dative, that can explain both polysemy (as in the case of the German dative, see § 9.4.5.3, which also has the functions of a locative within prepositional phrases) and merger, as in the case discussed here.

Mycenaean Greek had a separate instrumental, which later merged with the dative-locative: this is the situation attested in all other Greek sources, starting with the Homeric poems. Syncretism of the dative-locative with the instrumental can be explained if we take the locative meaning as starting point. As already remarked in § 9.2, some referents have the peculiarity that they can be conceptualized as instruments or as locations: among them, one can think of means of transportation or body parts (for example, even in a language like Russian, which has a separate instrumental case, one would commonly used the preposition *na* 'in' with means of transportation: *poexat' na poezde, na mašine* 'to go by train, by car, lit.: in train, in car').[6] Thus, the meaning of the Classical Greek dative constitutes a radial category, with the location schema at its centre.

[6] The tendency for body parts to be conceived as containers while used as instruments is especially clear in ancient Greek, see Luraghi (2004a).

9.5.2 The direction of semantic extension

It is not always easy to find enough diachronic evidence to say that a case has acquired a certain meaning later than another. In some Indo-European languages there is at least one clear instance of semantic extension of a case, i.e. the extension of the instrumental case to encode agent of passive verbs. The reason why we can be sure that the instrumental meaning is older, while the agentive meaning is later, is that comparative evidence points toward a late development of passive diathesis. Besides, the instrument of agent only exists in a small number of languages (Indo-Iranian and, partially, Slavic); in Hittite, a language in which early sources offer no evidence for the existence of passive diathesis, the instrumental of agent only occurs in comparatively late texts, while the instrumental of instrument is attested from the beginning.[7]

The semantic extension instrument > agent contradicts the (putatively unidirectional) scale of semantic extension posited in Heine et al. (1991: 159). According to the authors, there is a universal tendency for spatial relations to precede all other types of relation, followed by relations involving human beings, and then by relations involving inanimate entities, which can be summarized as:

> *spatial relation* > *human relation* > *inanimate relation*
> (from Luraghi 2001).

In Luraghi (2001) I have argued that evidence from Australian languages allows us to reconstruct the opposite direction of semantic extension, i.e. from agent (ergative case) to instrument, in accordance with the predictions in Heine et al. (1991). This extension can be motivated by a metaphor, according to which 'an instrument is an agent': the metaphor is based on the feature of ultimate control over the event.[8] Thus, the extension attested in the Indo-European languages remains to be explained. In Luraghi (2001) I tentatively suggested that this extension could be motivated by a metonymy, rather than a metaphor, but the matter needs further investigation.[9]

[7] For further discussion see Luraghi (1986) and (2001).

[8] Note that the metaphor involves personification, much in the same way as the well-known 'companion metaphor', discussed in Lakoff and Johnson (1980). As for the feature of control, shift from agent to instrument is made possible by a gestalt effect: obviously, only the agent can fully control an event, the instrument can exert control only if one abstracts from other features typical of agents, like intentionality. Based on a similar gestalt effect we can say things like *the key opened the door*, treating an instrument as if it were responsible for an action.

[9] In general the bearing of metonymy on semantic change seems to be underestimated in CG-oriented studies of the meaning of grammatical forms.

9.5.3 Relative markedness

In his critique of Lakoff and Johnson's 'companion metaphor', Stolz (1998) notes, among other things, that, whereas the meaning extension seems to only lead from comitative to instrumental, there is morphological evidence for derivation of comitative markers from instrumental ones. Stolz mentions the Australian language Alyawarra, which has an instrumental suffix *-ila*; and a comitative *-ila-linga*, derived from the instrumental by addition of another suffix. Such a state of affairs is also common in the Indo-European languages: for example, in Latin instrument is encoded by the ablative case, and comitative by the ablative with the preposition *cum*, in Russian instrument is encoded by the instrumental case, and comitative by the instrumental case with the preposition *s*, and so on.

The Indo-European languages, that allow far reaching diachronic analysis, show that there is no contradiction in the fact that the comitative is more marked than the instrumental, while cognitively preceding it. Indeed, there is evidence for the semantic extension from comitative to instrument, but apparently after the extension, comitative tends to receive extra marking, in much the same way as one often says 'together with' in English, rather than simply 'with' to denote comitative.[10]

9.5.4 From polysemy to homonymy?

Above I have illustrated the complex syncretism that led the Classical Greek dative case to also encode relations that are typical of the locative and the instrumental. I have argued that locative was at the centre of the category. Note that in this way, locative provided an otherwise missing link between instrument and dative. However the discussion above does not tell the whole story. The extent to which the dative could encode location was indeed very limited already in Homeric Greek; later, in Classical Greek prose, the dative without prepositions only encoded the functions of the dative 'proper' (recipient, beneficiary, experiencer, and the like), or the functions of the instrumental. The locative meaning was limited to the prepositional dative, and even with a number of prepositions the association of dative and locative had lost motivation (see Luraghi 2003 for the details). Now the question arises if, once the centre of a radial category is lost, its peripheral members are still recognized as belonging to the same category, in other words, whether we can still regard such cases as instances of polysemy, or should better regard them as (secondary) homonymy.

Perhaps the Greek dative case is not the best example in this respect, since, as I said, part of the locative meaning survived with prepositions; in any case, the dative disappeared and was variously substituted by prepositions (or by the genitive in

[10] I do not have space here to review the evidence, but see Luraghi (2001).

978–0–19–920647–6 10-Spencer-ch09 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 150 of 150 May 16, 2008 18:57

the case of pronouns), so the study of its semantic development cannot be pursued further. However, the question remains: if some historically attested semantic extension loses motivation, can it give rise to homonymy? In Luraghi (forthcoming) I have suggested that such a process can explain the development of the Greek preposition *metá*, which could mean 'with' (with the genitive) and 'after' (with the accusative) in Classical Greek. The two meanings can be shown to be related when they originated out of the original meaning 'among' in Homeric Greek, but the centre of the category was lost, and the two meanings became so disconnected that in Medieval Greek the preposition underwent two phonological developments: on the one hand, it continued the older form when it had the meaning 'after', on the other, it changed to *mé* when it had the meaning 'with'. The latter is the only form that survives in Modern Greek.

The above example deals with a preposition, rather than with a case, but similar evidence can most likely be found in the development of case markers as well. I think that the fact that we regard polysemy as the normal situation of grammatical forms (or better, of any linguistic item) should not necessarily imply that there is no possible development by which two meanings become so far from each other as to make a synchronic association impossible for speakers, i.e. as to generate homonymy. Indeed, I view it as a challenge of CG to set limits to the explanatory power of semantic extension: possible creation of homonymy looks very much like one of these limits.

CHAPTER 10

CASE IN NSM

A REANALYSIS OF THE POLISH DATIVE

ANNA WIERZBICKA

10.1 Introduction: A Boasian perspective on grammatical categories[1]

The NSM theory of language (see below, section 10.2) assumes, with Franz Boas, that the grammatical categories of a language represent a given speech community's interpretation and classification of experience, in a broad outline. As Boas put it, 'Since the total range of personal experience which language serves to express is infinitely varied, and its whole scope must be expressed by a limited number of phonetic groups, it is obvious that an extended classification of experiences must underlie all articulate speech' (Boas 1964[1911]: 121).

To many contemporary linguists, especially cognitive linguists, Boas' view of grammatical categories may seem so uncontroversial as to be hardly worth recalling. For Boas, however, this view had important methodological consequences

[1] This chapter owes a great deal to discussions with Cliff Goddard, both on the level of general ideas and on the level of detail. I am also grateful to an anonymous reviewer, whose comments and suggestions were very helpful in revising an earlier draft of the chapter.

which are at variance with most contemporary approaches, 'cognitivist' or otherwise.

From the perspective of the NSM theory of language, the most remarkable aspect of Boas' vision of grammar was his adoption of 'the native speaker's point of view': when he spoke of 'categories of thought' encoded in grammar he didn't mean some abstract and esoteric meanings expressible only in technical language, but meanings that intelligent native speakers could arrive at by analysing their own thoughts. Thus, in conclusion of the 'Introduction' to his *Handbook of American Indian Languages* he wrote that in his work 'the grammar has been treated as though an intelligent Indian was going to develop the forms of his own thoughts by an analysis of his own form of speech' (Boas 1964[1911]: 123). This is also the approach taken in the NSM research into the semantics of grammar (cf. e.g. Wierzbicka 1980, 1981, 1988; Goddard 2003), including the semantics of cases.

The problem with Boas' account, however, was (as NSM researchers see it) that he didn't have at his disposal a semantic metalanguage which would have allowed him to identify the 'groups of ideas' or 'categories of thought' chosen (unconsciously) by the speakers of a given language as a basis for specific grammatical categories (or, for that matter, lexical ones: words and phrasemes). It is this problem that the NSM theory, from its inception, has focused on and has sought to address (as its name, derived from the phrase 'the natural semantic metalanguage', indicates).

10.2 Metalanguage as a central problem for semantic analysis

Most contemporary approaches to semantics in general and the semantics of grammar in particular treat the issue of a suitable metalanguage as non-existent. The common practice is to use technical or semi-technical English as a metalanguage, with little or no justification and without any standardization.

For example, discussions of the meaning of cases often rely on English terms like *experiencer*, *recipient*, *benefactive*, *affected*, *affectee*, *goal*, *endpoint*, *transfer*, *purposive*, and the like, which are either defined ad hoc, or not defined at all, as if their meaning were self-explanatory. When definitions *are* offered, they tend to be formulated in words chosen ad hoc and not any more self-explanatory than those which they are purporting to explain. Not surprisingly, the resulting debates about what an 'agent' or an 'experiencer' 'really is' lead nowhere.

From an NSM point of view, the failure to address the issue of a suitable metalanguage is the sore point of most contemporary approaches to the semantics of cases (as well as other areas of lexical and grammatical semantics).

The most distinctive feature of the NSM approach is that it takes seriously the idea advanced by seventeenth-century European philosophers like Descartes, Arnauld, and above all Leibniz, that only a small repertoire of self-explanatory simple concepts can provide the bedrock of all human understanding. To quote Leibniz: 'If nothing could be understood in itself nothing at all could ever be understood. Because what can only be understood via something else can be understood only to the extent to which that other thing can be understood, and so on; accordingly, we can say that we have understood something only when we have broken it down into parts which can be understood in themselves' (Leibniz 1903/1961: 430).

The NSM approach to semantics has adopted this idea, and its practitioners have engaged, over more than three decades, in theoretical and empirical investigations, seeking to identify, by trial and error, a set of self-explanatory 'semantic primes' which could free semantic analysis from infinite regress and lead to genuine understanding of the meanings encoded in lexicon and in grammar. To this end, NSM researchers have undertaken wide-ranging experimentation over many semantic domains, across many diverse languages. The purpose of this experimentation was to identify within the languages under investigation matching minimal sets of lexically embodied simple meanings in terms of which all other, more complex meanings could be intelligibly explained and compared.

10.3 The 'natural semantic metalanguage' (NSM)

The natural semantic metalanguage, built through extensive cross-linguistic investigations, is described in great detail in various publications, such as Goddard (1998) and Wierzbicka (1996), and especially Goddard and Wierzbicka (2002a), which also contains six studies demonstrating that the posited semantic primes and their basic syntactic frames exist in a set of typologically and genetically diverse languages. The full NSM lexicon of universal semantic primes is set out, in summary form, in Table 10.1, using English and Polish exponents.

Detailed studies of a range of other languages have also been carried out over the past fifteen or so years, so that there is now a substantial body of cross-linguistic work on the realizations of semantic primes. Some of the languages and investigators include: Chinese (Chappell 2002), Mbula (Bugenhagen 2002), Amharic (Amberber, in press), Lao (Enfield 2002), Spanish (Travis 2002; Aznárez 2005), Korean (Yoon 2006, 2007), Ewe (Ameka 1994), French (cf. Peeters 2006), Polish (Wierzbicka 2002), East Cree (Junker 2003, in press; Junker and Blacksmith 2006), Japanese (Hasada 1997; Onishi 1994, 1997), Russian (Gladkova, forthcoming). A

Table 10.1. Semantic primes – English and Polish exponents

Substantives:	I, YOU, SOMEONE, SOMETHING/THING, PEOPLE, BODY	JA, TY, KTOŚ, COŚ, LUDZIE, CIAŁO
Relational substantives:	KIND, PART	RODZAJ, CZĘŚĆ
Determiners:	THIS, THE SAME, OTHER/ELSE	TEN, TEN SAM, INNY
Quantifiers:	ONE, TWO, MUCH/MANY, SOME, ALL	JEDEN, DWA, DUŻO, NIEKTÓRZY/NIEKTÓRE, WSZYSCY/WSZYSTKO
Evaluators:	GOOD, BAD	DOBRY, ZŁY
Descriptors:	BIG, SMALL	DUŻY, MAŁY
Mental predicates:	THINK, KNOW, WANT, FEEL, SEE, HEAR	MYŚLEĆ, WIEDZIEĆ/ZNAĆ, CHCIEĆ, CZUĆ, WIDZIEĆ, SŁYSZEĆ
Speech:	SAY, WORDS, TRUE	POWIEDZIEĆ/MÓWIĆ, SŁOWO, PRAWDA
Actions, events, movement, contact:	DO, HAPPEN, MOVE, TOUCH	ROBIĆ, DZIAĆ SIĘ/STAĆ SIĘ, RUSZAĆ SIĘ, DOTYKAĆ
Location, existence, possession, specification:	BE (SOMEWHERE), THERE IS/EXIST, HAVE, BE (SOMEONE/SOMETHING)	BYĆ (GDZIEŚ), BYĆ/ISTNIEĆ, MIEĆ, BYĆ (CZYMŚ/KIMŚ)
Life and death:	LIVE, DIE	ŻYĆ, UMRZEĆ
Time:	WHEN/TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT [IN ONE MOMENT]	KIEDY/CZAS, TERAZ, PRZED, PO, DŁUGO, KRÓTKO, PRZEZ PEWIEN CZAS, CHWILA [W JEDNEJCHWILI]
Space:	WHERE/PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE [ON ONE SIDE], INSIDE	GDZIE/MIEJSCE, TUTAJ, NAD, POD, WEWNĄTRZ/W, STRONA [Z JEDNEJ STRONY], BLISKO, DALEKO
Logical concepts:	NOT, MAYBE, CAN, BECAUSE, IF	BO /Z POWODU, JEŻELI, NIE, BYĆ MOŻE, MÓC
Augmentor, intensifier:	VERY, MORE	BARDZO, WIĘCEJ
Similarity:	LIKE	TAK/ TAK JAK

Notes: • Primes exist as the meanings of lexical units (not at the level of lexemes) • Exponents of primes may be words, bound morphemes, or phrasemes • They can be formally complex • They can have different morphosyntactic properties, including word-class, in different languages • They can have combinatorial variants (allolexes) • Each prime has well-specified syntactic (combinatorial) properties. • Two (or more) primes can share the same lexical exponent, with different syntactic properties.

great deal of descriptive semantic work has been conducted in these and other languages. A sizable bibliography (including references to the above) is available at: http://www.une.edu.au/lcl/nsm/index.php.

Empirical investigations carried out within the NSM framework suggest that similar tables can be drawn up for any language, and moreover, that the semantic primes listed in them exhibit in all languages the same combinatory properties. This means that every language has as its semantic core a language-like structure, with a mini-lexicon and a mini-grammar. Each such mini-language is in fact a surface realization of one and the same underlying system, both 'natural' and formal (cf. Lehrman 2006). Crucially, this universal 'mini-language' can be used effectively as a 'natural semantic metalanguage' for the description and comparison of both lexicons and grammars.

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 155 of 169 May 16, 2008 18:58

As this chapter seeks to illustrate, the use of this mini-language opens up an insider perspective on the meaning of grammatical categories. It circumvents circularity, obscurity, and terminological ethnocentrism and it makes the representations of meaning accessible to the speakers concerned. As this chapter also seeks to illustrate, empirically established universal semantic primes provide a *tertium comparationis* which is stable and language-neutral, is capable of representing the conceptual reality of native speakers, allows maximum resolution of semantic detail, ensures the possibility of empirical disconfirmation and provides a supple and precise tool for systematic cross-linguistic comparisons.

In this chapter, I will show how NSM can be applied to the study of cases by revisiting my earlier (1986a) study of the Polish dative. Since that earlier study the metalanguage has been significantly expanded and revised in the light of empirical cross-linguistic investigations (cf. in particular *Semantic and Lexical Universals*, Goddard and Wierzbicka 1994, and *Meaning and Universal Grammar* [M.U.G.], Goddard and Wierzbicka 2002a), and as this chapter hopes to show, its current, 'post M.U.G.'[2] form offers a more precise and more effective tool for exploring the meaning of cases.

At the same time, this chapter continues the semantic approach to cases which was launched, in opposition to the then prevailing 'autonomous syntax' approaches, in the author's 1980 book *The Case for Surface Case.*

10.4 Psychologically plausible cognitive scenarios for Polish cases

Looking at the Polish cases described traditionally as nominative, accusative, instrumental, and dative from a Boasian perspective (that is, from the perspective of an intelligent insider) one detects above all an interest in aspects of human action.

In one prototypical scenario, the speaker is talking simply about someone doing something. The case used to identify this someone is the nominative, for example:

Adam *pracuje.*
Adam.NOM work.PRES.3SG
'Adam is working.'

In the second prototypical scenario, the speaker is talking about someone who did something to something and who wanted something to happen to this thing. In

[2] The acronym M.U.G. stands for *Meaning and Universal Grammar* (Goddard and Wierzbicka 2002a), a landmark publication in NSM semantics. The form of NSM presented in that book, and in subsequent NSM publications, is frequently referred to as 'the new NSM'.

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 156 of 169 May 16, 2008 18:58

this case, the doer is still marked by the nominative, and the target object by the accusative. For example:

Adam *porąbał* *drwa.*
Adam.NOM chop.PAST.PRF.3SG wood.ACC.PL
'Adam chopped (up) the blocks of wood.'

In the third prototypical scenario, the speaker is talking about someone who was doing something to something with some other thing. In this case, the doer is marked by the nominative, the target object by the accusative, and the 'other thing', which is the instrument of the action rather than its target, by the instrumental. For example:

Adam *rąbał* *drwa* *toporem.*
Adam.NOM chop.PAST.IMPF.3SG wood.ACC.PL axe.INS
'Adam was chopping the blocks of wood with an axe.'

In the fourth prototypical scenario, the speaker is talking about someone who did something to something because he or she wanted something to happen to someone else (for example, he or she wanted this other person to have, or to see, the thing in question). For example:

Adam *dał* *coś* *Ewie.*
Adam.NOM give.PAST.3SG something.ACC.PL Eve.DAT
'Adam gave something to Eve.'

Obviously, not all uses of the four cases in Polish would fit the four scenarios outlined above: what is being suggested here, then, is not some 'general meanings' (along the lines of Jakobson's 'Gesamtbedeutungen') but certain prototypical cognitive scenarios on the basis of which the grammatical categories in question can be identified by linguists and which, arguably, constitute conceptual guideposts, or reference points, for the speakers themselves.

Using NSM, one can formulate the four prototypical scenarios discussed above as follows:

Scenario I (a prototype for the nominative)
- someone did something

Scenario II (a prototype for the accusative)
- someone did something to something
- because this someone wanted something to happen to this thing
- something happened to this thing because of it

Scenario III (a prototype for the instrumental)
- someone did something to something with something else
- because this someone wanted something to happen to this thing

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 157 of 169 May 16, 2008 18:58

Scenario IV (a prototype for the dative)
- someone did something
- because this someone wanted something to happen to someone else
- something happened to this other someone because of it

To see that NSM semantic representations are indeed not only intuitively intelligible (and testable) but also cross-translatable, readers familiar with Slavic languages can consider the Polish versions of the four prototypical scenarios:

Scenario I

I. ktoś coś zrobił
II. ktoś coś z czymś zrobił
ponieważ ten ktoś chciał, żeby coś się stało z tym czymś
coś się stało z tym czymś tego powodu
III. ktoś coś czymś zrobił z czymś innym
ponieważ ten ktoś chciał, żeby coś się stało z tym czymś innym
IV. ktoś coś zrobił
ponieważ ten ktoś chciał, żeby coś się stało komuś innemu
coś się stało temu komuś innemu z tego powodu

The morphological trappings in the Polish and English formulae are different, and so is the word order, and also the 'resonance' (cf. Goddard and Wierzbicka 1994), but the meaning (in the sense of paraphrasable meaning) is exactly the same.[3]

10.5 Polish transitive constructions with the dative

Polish constructions using the dative fall into two broad classes: transitive and intransitive ones. For reasons of space, only transitive constructions with a verb of 'doing' can be discussed here in NSM terms, and not even all of those. The focus here is not on exhaustiveness but on the methodology. The overall theme of these Polish transitive constructions (with the dative and a verb of 'doing') corresponds to scenario IV above. For example, the sentence:

Adam *zrobił* *Ewie* *zastrzyk.*
Adam.NOM did Eve.DAT injection.ACC
'Adam gave Eve an injection.'

[3] Paraphrasability is of course relative to a particular set of primes. What is meant here is there are no differences between the two versions which could be paraphrased in the set of primes specified in Table 10.1 – and this set has no competitors in the existing linguistic literature.

implies that Adam did something because he wanted something to happen to Eve and that something happened to Eve because of this.

Apart from lexically governed datives, which will not be included in my discussion, Polish transitive constructions with the dative and a verb of 'doing' can be divided into several classes on the basis of the specific effect that the doer's action has on the 'affected' person (i.e. the person designated by the dative). Five major classes can be characterized in NSM terms as follows (a 'wanting' component is shared by all these classes):

1. someone did something (to something) at some time
   someone else could **have** something because of this
2. someone did something (to something) at some time
   someone else could **do** something (to this something) because of this
3. someone did something (to something) at some time
   someone else could **see/hear/know** something because of this
4. someone did something (to something) at some time
   someone else could **feel something good/bad** because of this
5. someone did something (to something) at some time
   someone else could **feel something in his/her body** because of this

Below, I will present a more detailed account of these five major classes, with examples, explications, and brief commentaries. For reasons of space this account cannot of course be exhaustive.

10.5.1 'Have' datives

Examples

Sprzedała mu samochód. 'She sold him a (the) car.'
Przywiozła mu książki. 'She brought him [by vehicle] books.'
Posłała mu paczkę. 'She sent him a parcel.'
Rzuciła mu piłkę. 'She threw him the ball.'
Kupiła mu krawat. 'She bought him a tie.'
Zrobiła mu na drutach sweter. 'She knitted him a sweater.'

Explication

- someone did something at some time
- because this someone wanted something to happen to someone else
  (this someone wanted this other someone to have something)
- something happened to this other someone because of it
  (this other someone could have this something because of this)[4]

[4] In English, phrases like *could have* or *could do* are ambiguous between a past and a conditional reading. The ambiguity of the English word *could* does not point to any inherent polysemy of CAN,

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 159 of 169 May 16, 2008 18:58

Sentences falling into this broad class could be further subdivided on the basis of a number of additional criteria, in particular, on the basis of the possible syntactic alternations. Despite the differences, all these subtypes are compatible with the construal presented in the unitary explication formulated above. This construal allows not only for 'transfer of ownership' (as in *sprzedać* 'sell') or physical 'transfer' (as in *rzucić* 'throw') but also for other situations, such as 'making', 'sewing', or 'buying' something for someone. 'Transfer' is too blunt a tool to capture the semantic common denominator here. The NSM formula of 'wanting this other someone to have something' offers a sharper characterization.

The use of the dative does not guarantee that the 'other person' did have something as a result of the action (e.g. the parcel may have gone astray), but it implies more than a mere intention on the part of the doer, as a phrase with the preposition *dla* 'for' would. The NSM formula 'this other someone could have this thing because of this' is sufficiently vague to cover the whole range of the 'have' datives (cf. Bacz 1991, Rudzka-Ostyn 1996) while still differentiating this range from that of the competing prepositional phrases.

To see this, compare for example sentences A, B, and C below:

A. *Kupiła dla niego sweter.*
'She bought a sweater for him.'

B. *Kupiła mu sweter.*
'She bought him [DAT] a sweater.'

C. *Dała mu sweter.*
'She gave him [DAT] a sweater.'

Sentence A could be expanded as follows: *nie wiedząc, że on już nie żyje* ('not knowing that he had already died'), but B and C could not. Technical phrases like 'affected', 'affectee', or 'recipient' cannot explain these facts, but NSM formulae can. Thus, in all three cases, 'she' did something (bought a sweater) because she wanted 'him' to have a sweater. In B and C, 'he' could have had a sweater because of this (a meaning contributed by the dative). In C alone, it is clear that 'he' did in fact have a sweater because of this (information contributed by the lexical meaning of the verb *dać* 'give').

Neither B nor C are compatible with the expansion 'not knowing that he had already died' because the dative carries the component 'he could have it because of this' (he couldn't if he was already dead). Sentence A, however, does not include this component and so it is compatible with the addition.[5]

but to an idiosyncratic gap in English. In other languages, exponents of CAN can be readily combined with exponents of the past tense, without any resulting ambiguity. For example, the Polish phrase *on mógł mieć* 'he could [PAST] have' is appropriately vague, not ambiguous. (For discussion, see Goddard and Wierzbicka 2002b: 74.)

[5] The English verb *have* is of course highly polysemous. The meaning which corresponds to the NSM prime HAVE is that which has its converse in the English verb *belong*. 'Possession' as a linguistic term is used much more widely, and much more variably.

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 160 of 169 May 16, 2008 18:58

10.5.2 'Can do' datives

Examples

Piotr umył mu filiżankę.
'Peter washed him a cup (to drink from).'
Piotr zatemperował mu ołówek.
'Peter sharpened him a pencil (to write with).'
Piotr otworzył mu drzwi.
'Peter opened him the door'.

Explication

- someone did something to something at some time
- because this someone wanted something to happen to someone else (this someone wanted this other someone to be able to do something)
- something happened to this other someone at that time because of this (this other someone could do something because of this)

Many sentences of this kind can be translated into English by means of the 'internal dative' construction, but only those which imply some tangible effect on the object, for example (cf. Wierzbicka 1986b):

Piotr otworzył Pawłowi puszkę sardynek.
'Peter opened Paul a tin of sardines.' [for Paul to eat]
Piotr otworzył Pawłowi drzwi.
*'Peter opened Paul the door.'

In Polish, in contrast to English, this construction allows also sentences which do not imply any tangible effect on the object, as long as it is clear what the person designated by the dative can do as a result. For example, the action of closing the door (in contrast to opening) would normally not be described in Polish with the dative:

**Piotr zamknął Pawłowi drzwi.* 'Peter closed Paul the door.'

Technical labels like 'beneficiary', 'affected', or 'experiencer' cannot account for such differences between languages, because to the extent to which such labels are intelligible they would apply to English sentences like *Peter opened Paul a tin of sardines* as much as to **Peter opened Paul the door*, and so they would make false predictions. Using NSM, on the other hand, we can tease out the relevant semantic differences. In this case, it is the difference between 'wanting someone to be able to

Rappaport Hovav and Levin (2005) link the meaning of some 'dative verbs' (those 'inherently involving possession, such as English *have*, *own* or *possess*') with what they call 'primitive predicate HAVE', without explaining the relation of this 'primitive' to the NSM prime HAVE. Their example sentences, however, suggest that their notion of 'possession' is much more abstract than that of the NSM prime HAVE, and its exact meaning is not clear.

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 161 of 169 May 16, 2008 18:58

do something' (Polish) and 'wanting someone to be able to do something **to this thing**' (English). The Polish formula is broader and correctly predicts a wider range of use, including being able to pass through a door (without doing something to that door).

10.5.3 'See/hear/know' datives

Examples

Pokazała mu fotografię. 'She showed him the photo.'
Zagrała mu walca. 'She played him a waltz.'
Powiedziała mu prawdę. 'She told him the truth.'

Explications

A. *Pokazała mu fotografię*. 'She showed him the photo.'

- someone did something to something at some time
- because this someone wanted something to happen to someone else (this someone wanted this other someone to see something)
- something happened to this other someone at that time because of this (this other someone could see this thing because of this)

B. *Zagrała mu walca*. 'She played him a waltz.'

- someone did something at some time
- because this someone wanted something to happen to someone else (this someone wanted this other someone to hear (this) something)
- something happened to this other someone at that time because of this (this other someone could hear something because of this)

C. *Powiedziała mu prawdę*. 'She told him the truth.'

- someone said something to someone else at some time
- because this someone wanted something to happen to this other someone (this someone wanted this other someone to know something)
- something happened to this other someone at that time because of this (this other someone could know something because of this)

On a very abstract level these three explications based, respectively, on 'seeing', 'hearing', and 'knowing' could be reduced to one, along the lines of formula D below:

D.
- someone did something at some time
- because this someone wanted something to happen to someone else
- something happened to this other someone at that time because of this

But the predictive power of such a 'leaner' formula would be reduced, as not all sentences are compatible with the use of the dative. In particular, such a formula

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 162 of 169 May 16, 2008 18:58

would fit sentences like *zamknęła mu drzwi* 'she closed him the door', which as already noted are in fact not acceptable in Polish.

10.5.4 'Feel good/ bad' datives

The third class of dative constructions to be discussed here includes two 'beneficial' and two 'maleficial' ones. According to the folk philosophy reflected in these constructions, both the 'benefits' and the 'harms' can come to people in two ways: via their possessions and via their close relatives. Thus, a doer can affect another person – for the good or for the bad – either by doing something to one of this person's possessions or to one of his or her close relatives or quasi-relatives (in NSM, someone who is 'like a part' of this person). Roughly, four subtypes of the class under discussion can be described as follows:

1. Something good happens to one of someone's possessions.
2. Something bad happens to one of someone's possessions.
3. Something good happens to one of someone's close relatives.
4. Something bad happens to one of someone's close relatives.

Examples, explications, and brief comments on each of these four subtypes follow.

10.5.4.1 *Something good happens to one of someone's possessions*

Examples

Nareperowała mu zegarek.
'She repaired "him" [his] watch.'
Oczyściła mu buty.
'She cleaned "him" [his] shoes.'

Explication

- someone did something to something at some time
  because this someone wanted something good to happen to someone else
- something happened to this thing because of this
- because of this, after this, this thing was not like before
- this thing belonged to this other someone
- this other someone wanted this thing to be like this
- because of this, when this happened to this thing it was good for this other someone
  (this other someone could feel something good because of this)

Here, the doer's action is directed at something that belongs to someone else and it is intended to be beneficial for this other person. The action changes in some way the state of the object in question. Since the change corresponds to the owner's

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 163 of 169 May 16, 2008 18:58

wishes, it is beneficial to the owner and can be expected to result in some good feelings on his or her part.

10.5.4.2 *Something bad happens to one of someone's possessions*

Examples

Zepsuła mu zegarek.
'She wrecked "him" [his] watch.'
Poplamiła mu marynarkę.
'She soiled/dirtied "him" [his] jacket.'

Explication

- someone did something to something at some time
  because this someone wanted to do something
- something happened to this thing because of this
- because of this, after this, this thing was not like before
- this thing belonged to someone else
- this other someone wanted this thing to be not like this
- because of this, when this happened to this thing it was bad for this other someone
  (this other someone could feel something bad because of this)

In many respects, this explication is symmetrical to the previous one – but not in all respects. The main difference is that in the case of a beneficial action, it appears that the beneficial effect on the other person is intended, where in the case of a detrimental one, it can be not intended. Nonetheless, in the 'detrimental' case, too, the doer is blamed by the speaker and is seen as responsible for the result. In this case, however, no intention of causing harm for the other person is attributed to the doer, instead, the doer's action is presented by the speaker as deliberate.

The asymmetry between the two constructions (as analysed here) may be displeasing for a semanticist for whom simplicity and elegance of the analysis is of paramount importance. From a more human point of view, however, the asymmetry makes sense: the folk philosophy apparently reflected in these constructions implies that it usually takes some effort and some good will to benefit another person, whereas to damage something it is enough to be careless or self-absorbed (one can't repair another person's watch accidentally, but one can certainly wreck it accidentally, through 'culpable carelessness').

10.5.4.3 *Something good happens to one of someone's close relatives*

Examples

Uratowali mu syna.
'They saved "him" [his] son.'
Wyleczyli mu żonę.
'They healed "him" [his] wife.'

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 164 of 169 May 16, 2008 18:58

Explication

- some people did something to someone at some time
  because they wanted something good to happen to this someone
- something good happened to this someone because of this
- this someone was like a part of someone else
- because of this, when this happened to this someone
  this was good for this other someone
- this other someone could feel something very good because of this

By and large, this category parallels that in 5.4.1 ('She repaired "him" [his] watch.'), the main difference being that between benefiting someone by doing something to their property and benefiting someone by doing something to a person closely related to them (not just in the sense of kinship but emotionally: a person who was 'like a part of this someone'). Similar differences obtain also in the case of detrimental actions to which I now turn.

10.5.4.4 *Something very bad happens to one of someone's close relatives*[6]

Examples

Zabili mu żonę.
'They killed "him" [his] wife.'
Uwiódł mu córkę.
'He seduced "him" [his] daughter.'
**Torturowali mu żonę.*
'They tortured "him" [his] wife.'

Explication

- some people did something to someone at some time
  because they wanted something to happen to this someone
- something bad happened to this someone because of this
- because of this, after this, this someone was not like before
- this someone was like a part of someone else
- because of this, when this very bad thing happened to this someone
  this was bad for this other someone
- this other someone could feel something very bad because of this

As in the case of detrimental action to someone's property, in the present case, too, no intention is implied to hurt the undergoer's relative. What *is* implied is that

[6] The differentiation between 'bad/good' in section 10.5.4.3 and 'very bad' in 10.5.4.4 is deliberate: the construction described in 10.5.4.3 can be used for 'something bad' or 'something good', whereas the one in 10.5.4.4 can only be used for 'something very bad' (no counterpart with 'something bad' or with 'something very good').

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 165 of 169 May 16, 2008 18:58

the action produces a specifiable change of state in the undergoer. This explains why *zabić* 'to kill' can occur in this construction whereas *torturować* 'to torture' cannot (if X kills Y, afterwards Y is dead, but if X tortures Y, there is no specifiable subsequent change of state in Y).

10.5.5 'Body part' datives

Like many other languages, Polish has some special constructions for using the dative in relation to the human body. The conditions on such datives are often highly language-specific and as sections 10.5.5.1–3 illustrate, a word like 'affectee' is far too blunt a tool to capture the cross-linguistic differences in the conceptualization.

10.5.5.1 *Datives of 'invaded personal sphere': Polish vs. French*

If someone comes to be physically close to someone else, can this be regarded as 'affecting' this other person? English grammar makes no provision for such an idea, but French grammar does, because as noted by Bally (1926), sentences in the following pairs do not mean exactly the same:

a. *Il court derrière elle.*
'He is running behind her'.
b. *Il lui court derrière.*
'He "to-her" is running behind.'

a. *Il tourne autour d'elle.*
'He turns around her.'
b. *Il lui tourne autour.*
'He "to-her" turns around.'

a. *Il passe devant elle.*
'He passes in front of her.'
b. *Il lui passe devant.*
'He "to-her" passes in front of.'

While the first sentence in each pair is a mere statement of fact, the second presents the fact of someone's physical closeness as affecting the person whose 'personal space' has been entered:

- someone (X) is doing something near someone else's (Z's) body

Il lui tourne autour.=
- someone (X) is doing something near someone else's (Y's) body
- something is happening to this other someone (Y) because of this
- (this other someone (Y) can feel something because of this)

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 166 of 169 May 16, 2008 18:58

By contrast, in Polish someone's physical closeness can be considered as affecting another person only on condition that it involves specifically some parts of this other person's body, not just the body as a whole. Thus, one can say:

Ciągle stoisz mi za plecami.
'You are constantly standing "to-me" behind my back.'
Ciągle kręcisz mi się pod nogami.
'You are constantly moving "to-me" around [my] feet.'
Mucha lata ci koło nosa.
'A fly is flying "to-you" close to [your] nose.'

but not:

**Siedział mi naprzeciwko.*
'He was sitting "to-me" vis-à-vis.'
**Biegł mi obok.*
'He was running "to-me" next.'
*Kręcił mi się dookoła.
'He was turning "to-me" around.'

Thus, the concept grammaticalized in the Polish construction calls for a slightly different explication:

- someone (X) is doing something near some parts of someone else's (Y's) body
- something is happening to this other someone (Y) because of this
- (this other someone (Y) can feel something because of this)

10.5.5.2 *Looking at a place inside a part of someone else's body*

Examples

Zajrzała mu do gardła (nosa, ucha . . .)
'She looked "him" into [his] throat (nose, ear . . .).'

Explication

- someone did something at some time
- because this someone wanted to see a place inside some part of someone else's body
- something happened to this other someone because of this
- (this other someone could feel something because of this)

This subtype – which, again, has no equivalent in English – may seem to be overly specific, but any attempt to make it more general appears to lead to false predictions. In particular, looking briefly at a part of someone's body does not warrant the use of the dative in Polish anymore than in English. In the case of looking *into* a body part, however, the act can be seen as so intrusive that even a brief look

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 167 of 169 May 16, 2008 18:58

allows, indeed calls for, the use of the dative. The following two sentences, both with a perfective verb and with momentary implications, show the relevance of this difference to Polish grammar:

Zajrzała mu do ucha.
'She looked "him" [briefly] into [his] ear.'
**Spojrzał jej na nogi.*
'He looked "her" [briefly] at [her] legs.'

10.5.5.3 *'Staring' at some parts of someone's body: Polish vs. Italian*

Example

Patrzył jej na nogi (piersi).
'He was looking "her" at [her] legs (breasts).'

Explication

- someone was doing something for some time
- because this someone wanted to see some parts of someone else's body for some time
- something happened to this other someone because of this
- because this other someone could feel something because of this

Thus, in Polish even *looking at* some parts of another person's body can be seen as 'affecting' this other person – provided that one is doing it *for some time* (unlike in the case of *looking into*, where the condition of duration does not apply). On the other hand, as I noted elsewhere (Wierzbicka 1979, 1988), Polish does not extend the assumption of 'affectedness' as far as, for example, Italian, where not only prolonged *looking at* but even *seeing* can warrant the use of the dative:

**Widział jej majtki.* (Polish)
'He saw "her" [her] underpants.'
Le vedeva le mutandine. (Italian)
'He saw "her" [her] underpants.'

This is accounted for by a reference to 'doing something' in the explication above.

10.6 Concluding remarks

The hallmark of the NSM approach to the semantics of cases (or indeed any other area of either lexicon or grammar) is the use of the natural semantic metalanguage based on simple words with exact semantic equivalents in all languages.

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 168 of 169 May 16, 2008 18:58

One way of thinking about NSM is to treat it as a technique which can usefully supplement other, more conventional techniques of semantic analysis. To learn this technique one has to learn, above all, a strictly controlled subset of English, with matching subsets in other languages, and in particular, in the language under investigation.

The NSM methodology requires that explanatory formulae devised as semantic hypotheses be couched in simple non-technical words (such as DO, HAPPEN, SOMEONE, SOMETHING, WANT, and BECAUSE) – intelligible to native speakers and cross-translatable into other languages. It also requires that those words be combined in simple sentences, again, intelligible to native speakers and cross-translatable into other languages – sentences such as 'someone did something', 'someone did something to something with something else', 'someone wanted something to happen to someone else', 'someone wanted someone else to have something', and so on. (For detailed description of syntactic patterns which are cross-translatable and therefore allowed in NSM, see Goddard and Wierzbicka 2002a).

As the examples given above illustrate, the technique proposed here for the description of the meaning of cases relies, for practical reasons, on 'NSM English' (rather than, for example, 'NSM Polish'). This is a conscious choice: given the role of English in the contemporary world, including contemporary science and scholarship, it is inevitable that for the foreseeable future it is the English version of NSM, rather than the Polish, Chinese, or any other one, which can be proposed, realistically, as a practical metalanguage for cross-linguistic semantics and typology. This means, in a sense, giving the English version of NSM a privileged position as a metalanguage for linguistics. At the same time, however, replacing full-blown English with 'NSM English' (that is, with a reduced version of English isomorphic with all the other versions of NSM) means radically reducing (if not totally eradicating) the anglocentrism which currently pervades contemporary science, including linguistics, and which remains unrecognized and unacknowledged.

Do and *happen* are indeed English words, but so are *experiencer* and *affectee.* The two crucial differences are, first that *do* and *happen* have semantic equivalents in all other languages of the world, whereas *experiencer* and *affectee* do not; and second, that *do* and *happen* are English words intelligible to ordinary speakers of English, whereas *experiencer* and *affectee* are not.

As this chapter seeks to demonstrate, 'NSM English' – with its simple words like *do* and *happen* and their highly restricted syntax – provides an effective tool for describing and comparing meanings encoded in cases. Using this tool, we can arrive at more fine-grained analyses than those relying on technical English and we are able to capture subtle differences which can be neither detected nor explained by means of technical linguistic terminology (English or any other).

Forty years ago, shortly before his untimely death, Uriel Weinreich (1980: 389) argued that the main goal of semantics was to provide 'explicit, verbalized, validated meaning-descriptions' and that the key task was to 'search for a satisfactory

978–0–19–920647–6 11-Spencer-ch10 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 169 of 169 May 16, 2008 18:58

methodology for obtaining and validating such meaning descriptions'. He also argued that such descriptions should be intelligible to ordinary native speakers ('lay speakers') and should be capable of being validated in consultation with them. (For discussion, see Wierzbicka, in press.)

The NSM approach to semantics in general and the semantics of cases in particular offers such a methodology. NSM explications of case meanings such as those presented in this chapter are intended as 'explicit, verbalized, validated meaning descriptions' which can indeed be further tested, and if need be, amended, in consultation with 'lay speakers' whose conceptualizations they are intended to portray. Explications of this kind can serve practical goals like language teaching and language learning (cf. Goddard and Wierzbicka, in press), as well as facilitating genuine understanding of ways in which speakers of different languages interpret the world and categorize their experience.

CHAPTER 11

CASE IN FORMAL SEMANTICS

HELEN DE HOOP
JOOST ZWARTS

11.1 Introduction

Formal semantics is an approach to natural language semantics that uses logical and mathematical tools to model meaning. Since its introduction in linguistics in the early seventies it has been applied to a wide variety of phenomena. However, case has not received a lot of attention from formal semanticists, probably because the approach has mostly focused on languages with relatively sparse case systems. In this chapter, we will discuss the application of formal semantics to case in the domains of argument structure and space.

11.2 Case and argument semantics

In this section we will discuss work in which the central notions of grammatical function and noun phrase interpretation play an important role in relation to case-marking. We will discuss Keenan's (1989) Semantic Case Theory, and de Hoop's

978–0–19–920647–6 12-Spencer-ch11 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 171 of 182 May 16, 2008 18:58

(1992) and van Geenhoven's (1996) type-shifting approaches to case and voice alternations. One very influential line of research in formal semantics employs the mereological approach that models the nominal and verbal domain of semantic values in terms of algebraic structures. Following Krifka (1992) and Kiparsky (1998) we will show how this approach can account for the semantics of partitive case in Finnish.

11.2.1 Semantic Case Theory

Dowty (1982) and Keenan (1989) are two similar theories of the logical relations subject and object (that is, the first and second argument of a binary relation, e.g. a transitive verb) developed in the spirit of Montague grammar (Montague 1973). Both appeal to an abstract notion of 'case'. They take the abstract notions of 'subject' and 'object' corresponding to the first and second argument of a two-place predicate and extend the basic NP denotations with this extra meaning.

As Keenan (1989) points out, in principle transitive sentences are underdetermined as to which NP is interpreted as the subject and which as the object. Keenan provides a way of specifying this information without commitment as to how any particular language structures transitive sentences, or what case system is used, if any. He does so by expressing the argument structure of a sentence in terms of case extensions: 'nominative' (subject) and 'accusative' (object) case extensions are added to the basic denotations of the NPs. Given the denotation of a basic NP, for instance *Robert*, and a binary relation *kiss*, the nominative extension of *Robert* sends the binary relation *kiss* to the set of entities which Robert kissed. Similarly, the accusative extension of *Robert* sends *kiss* to the set of entities which kissed Robert:

(1) For F basic, F_{nom} or the *nominative case extension* of F is that extension of F which sends each binary relation R to {b: $F(R^b) = 1$} with $R^b =_{df}$ {a: (a,b) $\in$R}

(2) For F basic, F_{acc} or the *accusative case extension* of F is that extension of F which sends each binary relation R to {b: $F(R^b) = 1$} with $R^b =_{df}$ {a: (b,a) $\in$R}

Transitive sentences with two independent NPs are in principle four-ways ambiguous, but usually, however, languages constrain the possible interpretations. The independent NPs in a transitive sentence are interpreted by distinct case extensions. In English, word order determines that the first NP is interpreted 'nominatively' and the second one 'accusatively'. In Japanese, if there is an NP with suffix *-o*, it must be interpreted accusatively, otherwise a *-ga* suffixed NP is interpreted accusatively. This constraint accounts for the fact that (3) is not ambiguous, while (4) is.

(3) *Taroo-ga Hanako-o nagutta.*
Taroo Hanako hit
'Taroo hits Hanako.'

978–0–19–920647–6 12-Spencer-ch11 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 172 of 182 May 16, 2008 18:58

(4) *Taroo-ga Hanako-ga suki da.*
Taroo Hanako likes
'Taro likes Hanako.' or 'Hanako likes Taroo.'

In Warlpiri, if there is an NP marked with (the ergative case-marker) *-ngku*, it is interpreted nominatively, and otherwise a zero-marked NP is interpreted nominatively.

Dowty (1982) raises the question of how to deal with certain cases where there seems to be overlap between the roles of arguments and of modifiers. He discusses three cases in particular: indirect objects, the genitive alternation between *Jane's father* and *father of Jane*, and the agent phrase of a passive, expressed in a *by*-phrase. He suggests that the number of grammatical roles may greatly exceed the number of distinct morphological case forms employed by a language; 'hence there should be no embarrassment in principle to postulating two or more distinct rules which use the same case marking for NPs' (Dowty 1982: 118). The approaches by Keenan and Dowty do not try to find the 'meaning' of different cases at all. Both theories completely ignore differences between one type of structural case, such as accusative, and another, such as ergative, although there is some discussion on datives. Their notion of case is a highly abstract one.

To sum up, the only difference between Keenan's 'nominative' and 'accusative' case is the difference between the denotation of the first and the second NP argument of a transitive verb. In semantic type theory the two basic semantic types are *e* (for 'entity') and *t* (for truth value) and if *a* is a type and *b* is a type, then $<ab>$ is a type as well. The type $<ab>$ is a function that takes an argument of type *a* and yields an expression of type *b*. In Keenan's approach, transitive verbs denote in type $<e<et>>$, and they combine with deictic NPs (type *e*) or basic NPs (type $<<et>t>$). The latter type is also called 'quantificational' but it is the type that is in principle available for all NPs, whence Keenan's term 'basic'. Since Keenan interprets transitive verbs (binary relations) invariably as functions from the set of individuals to functions from the set of individuals to the set of truth values (type $<e<et>>$), the basic semantic types of 'nominative' and 'accusative' NPs do not change (both are either *e* or $<<et>t>$); only their extension (being interpreted as the subject or the object) differs. In the next section, we will discuss several formal semantic approaches to case-marking alternations that have abandoned this strict view, and propose more flexibility in the semantic composition of transitive sentences.

11.2.2 Case and type-shifting

Cross-linguistically, the semantic type of an NP seems to influence its case-marking. Let us illustrate this with an example of differential object marking. In Finnish, depending on the reading associated with it, the object of a transitive verb will be marked either with accusative or with partitive case:

978–0–19–920647–6 12-Spencer-ch11 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 173 of 182 May 16, 2008 18:58

(5) *Anne tapaa vieraita.*
Anne meets guests.PART
'Anne meets some guests.'

(6) *Anne tapaa vieraat.*
Anne meets guests.ACC
'Anne meets the guests.'

The partitive case in (5) triggers a non-specific (indefinite) reading, while the accusative case in (6) goes with a specific (definite) reading of the object. In traditional grammars (see e.g. Karlsson 1983), the alternation between a partitive object and an accusative object is attributed to two semantic distinctions, namely indefiniteness versus definiteness of the NP and irresultativity versus resultativity of the VP. This is exemplified by the alternation in (7)–(8):

(7) *Presidentti ampui lintua.*
President shot bird.PART
'The president shot at a/the bird.'

(8) *Presidentti ampui linnun.*
President shot bird.ACC
'The president shot a/the bird.'

While in (8) we get the resultative reading and the result of the shooting is clear (the bird was shot), in (7) the result is not known. That is, (7) is about the shooting rather than about the bird.

De Hoop (1989, 1992) proposes to unify these two semantic distinctions (NP-related and VP-related) by the notion of 'predicate modification'. Her basic idea is that the NP that bears partitive case can syntactically still be a 'full-fledged' argument (a 'real' object), but semantically it functions as 'part of the predicate', both in a context such as in (5)–(6) above where the NP differs in specificity, and in the other type of context, exemplified in (7)–(8) above, where the VP differs in resultativity.

Partee (1986) recognizes the existence of different semantic types of NPs and proposes type-shifting operations to deal with them. General type-shifting principles are responsible for going from one interpretation to the other in a certain context, thus allowing for more flexibility in relating syntax and semantics than Montague's (1973) strict correspondence between syntactic category and semantic type. As for NPs, besides the well-known referential type e, quantificational type $\ll et>t>$, and predicative type $<et>$, de Hoop proposes a fourth semantic type of NP, especially for languages with differential case-marking, namely the predicate modifier type $\ll et><et\gg$. She argues that object NPs that bear partitive case in Finnish are interpreted as 'part of the predicate'. Her idea is that syntactic case-marking in differential object-marking reflects type-shifting operations as proposed by Partee.

978–0–19–920647–6 12-Spencer-ch11 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 174 of 182 May 16, 2008 18:58

Transitive verb phrases can be composed in different ways. The verb can be truly transitive in the sense that it denotes a relation between two 'real' arguments (type *e* or type $\ll et>t>$), or the verb has the syntactic or semantic type of an intransitive verb (type $<et>$) with its object functioning as a predicate modifier (type $\ll et><et\gg$). Even a quantificational object bearing partitive case would be interpreted as a predicate modifier of type $\ll et><et\gg$ in the following example (de Hoop 1992):

(9) *Presidentti ampui kaikkia lintuja.*
president shot all.PART birds.PART
'The president shot at all birds.'

Within the same spirit, Ramchand (1993) shows that Scottish Gaelic behaves like Finnish with respect to case-marking and object interpretation. In the Finnish example in (9), a quantificational NP may appear with partitive case (in Scottish Gaelic this would be genitive case), and the reading that is obtained (both in Finnish and in Scottish Gaelic) is an irresultative (or unbounded) reading of the VP. As Ramchand puts it, a strong object plus irresultativity is marked the same way as a weak object. Ramchand argues that what this class of interpretational possibilities has in common is the particular relation between verb and object, that she also characterizes as predicate modification, following de Hoop (1992). Alternatively, Kiparsky (1998) argues that the common factor of the aspectual and object-related functions of Finnish partitive case is marking a VP's unboundedness (see the next subsection).

As pointed out in Bary and de Swart (2005) and de Swart (in prep.), the object in a cognate object construction behaves semantically as a manner adverb and should therefore also be treated as a predicate modifier of type $\ll et><et\gg$. In a relatively rich case-marking language like Russian this semantic type of the cognate object is mirrored by the use of an instrumental case in (10). In English, however, as the translation of Russian (10) shows, the cognate object is realized as any other direct object thus not showing its different semantic status (Pereltsvaig 1999).

(10) *On ulybnulsja ulybkoj angela.*
he.NOM smiled smile.INS angel.GEN
'He smiled an angelic smile.'

So far we have seen that transitive verb phrases can semantically be composed in different ways. The verb can be straightforwardly transitive (type $<e<et\gg$) in the sense that it denotes a relation between two arguments of type *e* (or $\ll et>t>$), or the verb can be formally intransitive (type $<et>$) with its object functioning as a predicate modifier (type $\ll et><et\gg$). A third option is that the verb functions as a predicate modifier (type $\ll et><et\gg$) which semantically incorporates a weak (predicative) object of type $<et>$, as proposed by van Geenhoven (1996). In West-Greenlandic, the incorporating verb is intransitive from a morphological

perspective. The incorporated noun can be modified by means of an adjective that bears instrumental case:

(11) *Esta nutaa-mik aalisagar-si-v-u-q.*
Ester.ABS fresh-INS.SG fish-get-IND-[TR]-3SG
'Ester got (a) fresh fish.'

The object in a semantic incorporation construction is not fully independent, but receives its existential quantificational force from the verb (cf. Carlson 1977). According to Van Geenhoven, a semantically incorporated object does not have to be morphologically incorporated as well, but it can also be realized as an NP bearing weak case. Certain verbs are inherently semantically incorporating (type $\ll et>\!<et\gg$), whereas others can never semantically incorporate their object.

Similarly, Partee and Borschev (2004) hypothesize that where we see nominative/genitive and accusative/genitive alternations in Russian (both under negation and under intensional verbs), the nominative or accusative represents an ordinary *e*-type argument position (or the basic NP type $\ll et>t>$), whereas a genitive NP is always interpreted as a property of type $<et>$. In the case of intensional verbs like *ždat'* 'expect, wait for' they argue that there is a shift in verb sense correlated with the shift in the interpretation of the object NP (Neidle 1988):

(12) *On ždet podrugu.*
He waits girlfriend.ACC
'He's waiting for his girlfriend.'

(13) *On ždet otveta na vopros.*
He waits answer.GEN to question
'He's waiting for an answer to the question.'

The object in (12) is specific, but the intensional variant of the verb *wait for* in (13) selects a non-specific NP. This would be a semantically incorporating type of verb in Van Geenhoven's approach, and the genitive object gets the concomitant predicative or property-denoting type. Partee and Borschev put forward the hypothesis that genitive subjects are property-denoting as well.

To sum up, the approaches discussed in this subsection were all attempts to account for morphological case alternations in terms of variation in the possible semantic types of NPs and verbs. Note that the result of the different options of semantic composition of a transitive verb and an object addressed in this subsection will in all cases be a VP of type $<et>$. Still, just as in the approaches of Keenan (1989) and Dowty (1982) that we have seen in the previous subsections, no attempt is being made to actually investigate the 'meaning' of a particular type of case. In the next subsection, we will discuss Kiparsky's (1998) study of the meaning of partitive case in Finnish. Following Krifka (1992), Kiparsky (1998) proposes that partitive case on an object NP in Finnish corresponds to the 'unbounded' meaning at VP-level.

978–0–19–920647–6 12-Spencer-ch11 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 176 of 182 May 16, 2008 18:58

11.2.3 Case and aspect

Krifka (1992) develops a theory that handles the influence of the reference type of NPs (mass, count) on the temporal constitution of VPs (activities, accomplishments). Nominal reference and temporal constitution are semantically related. For example, a quantized NP like *an apple* denotes an object with precise limits, just as *run a mile* denotes a telic event with precise limits. Similarly, a cumulative NP like *wine* denotes something without clear limitation, just like an atelic VP such as *run* has no clear limitation. The quantized object yields a telic (or bounded) VP in (14), while the cumulative object yields an atelic VP in (15):

(14) Jane drank a glass of wine in an hour.

(15) Jane drank wine for an hour.

Krifka explains how partitive case in Finnish can serve to mark progressive (irresultative) aspect. Krifka considers progressive predicates like *be drinking a glass of wine* as applying to events which are parts of events to which *drink a glass of wine* applies. On the other hand, partitivity can be associated with a similar operator: the partitive of *fish* can be analysed as referring to parts of a fish. A VP with a partitive object can denote an event which is a part of an event denoted by the corresponding VP with an accusative object. Kiparsky (1998) unifies the VP-related and NP-related functions of the Finnish partitive case in terms of boundedness of the VP. His generalization is that the object of an unbounded VP is obligatorily partitive. A VP is unbounded if it has either an unbounded verb or an unbounded object. Thus, the VP can be unbounded either in virtue of its head V or in virtue of its nominal object. But the morphological marking of this unboundedness is uniformly on the object NP in Finnish.

Kiparsky (1998) defines quantitatively indeterminate plurals and mass nouns as unbounded. Quantitatively indeterminate count nouns and cardinal noun phrases are bounded. Verbs such as *ostaa* 'buy', *tappaa* 'kill', *ottaa* 'take', and *mainita* 'mention', are bounded and assign accusative case to their objects. But when their objects are quantitatively indeterminate, they get the NP-related partitive case, and the VP as a whole becomes unbounded. Verbs that denote psychological states such as *halveksia* 'despise', *ihailla* 'admire', intensional verbs such as *pyytää* 'ask for', *ajatella* 'think about', and verbs that denote continuous motion or contact, like *ravistaa* 'shake', *koskettaa* 'touch', are unbounded and assign partitive case to all their objects. Finally, there are verbs which take both partitive and accusative objects depending on the VP's boundedness. Note that an event of painting a house may be unbounded (with the object *house* in the partitive), even if the whole house was in fact painted. But because the accusative object would imply that the whole house was painted, the use of the partitive carries a (defeasible) implicature that only part of the house was painted.

978–0–19–920647–6 12-Spencer-ch11 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 177 of 182 May 16, 2008 18:58

Partitive case on an object in Finnish thus reflects the unboundedness of the VP (not the unboundedness of the NP or the unboundedness of the verb per se). An accusative object can result in a bounded meaning for a VP. For example, *potkaista* 'kick' takes a partitive object in the meaning 'kick at', and an accusative object when a directional complement is added. Boundedness can be licensed by resultativity through explicit or implicit locative or resultative predication. In (17) the inherently unbounded verb *love* is construed as resultative:

(16) *Rakastin teitä.*
I.loved you (pl).PART
'I loved you.'

(17) *Rakastin teidät rappiolle.*
I.loved you(pl).ACC ruin.ADESS
'I loved you into ruin.'

To sum up, in Finnish partitive case licenses unboundedness at the VP-level, irrespective of where the unboundedness comes from (verb or object).

Kiparsky (1998) presents some further historical and comparative evidence suggesting that the partitive's emergence as a structural case is a precondition for the rise of its function to denote unboundedness at the VP-level. The first stage in the evolution of the Balto-Finnic partitive object was the partitive's change in status from a strictly locative case (meaning 'from') to case with quantificational force. Since locative cases must be interpreted in a local, compositional fashion, the grammaticalization of partitive case into a structural case in Finnish made it possible that it developed as a marker of unboundedness at the VP-level. The next section will deal with formal semantic analyses of spatial (locative and directional) case.

11.3 Case and spatial structure

The analysis of local cases is an area that can also benefit from the application of precise spatial models, as Fong (1997) and Kracht (2002) have shown in their analyses of the locative and directional dimensions of rich local case systems.

Languages with rich case systems are well known for their special *local* cases, like the inessive or illative in Finnish (18):

(18) a. *talo-ssa*
house-INESS
'in the house'
b. *talo-on*
house-ILL
'into the house'

978–0–19–920647–6 12-Spencer-ch11 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 178 of 182 May 16, 2008 18:58

Semantically, local cases function very much like adpositions (i.e. prepositions or postpositions). For example, if we interpret the English preposition *in* as a function from objects to their spatial interior, then the Finnish inessive will have to denote a very similar function. So, from a spatial point of view, there does not seem to be much to say about local cases beyond what we already know from adpositions, which might be the reason for the restricted attention that local cases have received in the semantic literature, either with a formal or cognitive orientation.

Nevertheless, there are at least a few important characteristics of local case systems that make them worthy of a semanticist's attention. First, because of their morphological and paradigmatic nature, local cases tend to encode spatial distinctions in a more systematic way than prepositions, especially in languages that are renowned for the richness of their case systems (Comrie and Polinsky 1998). Hence, analysing local case systems might reveal something about fundamental distinctions and patterns in spatial language. Second, even though local cases might be analogous to adpositions in their *semantics*, they are clearly part of the *morphosyntax* of noun phrases, with implications for the relation between syntax and semantics.

Before turning to the first topic, we should clarify an ambiguity in the term *local*, and related terms like *locative* and *locational*. Some authors use these terms in a general sense, for all adpositions or cases that are used with a spatial meaning, either involving places (where something is) or paths (where something is going). Others have used it in a more specific sense, in contrast to *directional*, to refer to expressions that refer to a place or location. In the wide sense, both expressions in (18) are local (spatial), but within this wide sense, we can say that (18a) is locational (place-referring), while (18b) is directional (path-referring).

In spite of the variety of labels used in the description of local cases, two components are usually distinguished, as illustrated in Table 11.1.

The vertical component is a particular spatial region relative to the noun (e.g. 'under'). This is what Jackendoff (1983) calls *place*, Comrie and Polinsky (1998) *orientation*, van Riemsdijk and Huybregts (2001) *location*, and Kracht (2002) *configuration*. The horizontal component is a type of motion with respect to that region (e.g. 'from'). Jackendoff calls this *path*, Comrie and Polinsky and Van Riemsdijk and Huybregts use the term *direction*, and Kracht *mode*. We will use Kracht's terms mode and configuration here.

Mode and configuration can be conflated into one portmanteau case morpheme, as in the Finnish illative *-on* (mode 'to' + configuration 'in'), but they can also be separately expressed, which is what we see in Lezgian postelative (Haspelmath 1993a):

(19) *sew-re-k-aj*
bear-AUG-under-from
from under the bear

Table 11.1. The two dimensions of local cases

	location –essive	source –elative	goal –lative, –directive	perlative, translative	versative, approximative
general	(ad)essive, locative 'at'	ablative, adelative 'from'	allative, addirective 'to'	'via'	'towards'
ad-					
in-	inessive 'in'	(in)elative 'out of'	illative, indirective 'into'	'through'	
super-	superessive 'on'	superelative 'off'	superdirective, superlative 'onto'		
sub-	subessive 'under'	subelative 'from under'	subdirective, sublative 'to under'		
post-	postessive 'behind'	postelative 'from behind'	postdirective 'to behind'		

In such cases we find the important hierarchical condition that the mode morpheme is outside (i.e. further away from the noun root than) the configuration morpheme. In the formal semantic analysis of Kracht (2002) mode and configuration are functions that together define a spatial relation between a moving or stationary object x and a reference object y over an interval I:

(20) mode (x, configuration (y), I)

A configuration is a function that, applied to an object y, gives a *parametrized neighbourhood*, a time-dependent set of spatial regions in or around y. Let us take a look at the 'in' configuration to see how the representation of $\text{in}'(y)$ is built up:

(21) $y \rightarrow \text{loc}'(y)(t) \rightarrow \iota(\text{loc}'(y)(t)) \rightarrow \{ r : r \subseteq \iota(\text{loc}'(y)(t)) \}$
$\rightarrow \lambda t \{ r : r \subseteq \iota(\text{loc}'(y)(t)) \}$

We start with the reference object y and derive its spatial location at time t, $\text{loc}'(y)(t)$, then define the interior of this region, extract all the regions included in this interior, and finally add the λ-operator to make it a function over times. This parametrized neighbourhood $\text{in}'(y)$ of y can then be the input for one of the five mode functions that Kracht distinguishes (Table 11.2).

The cofinal mode, for instance, builds on $\text{in}'(y)$ to define the following relation:

(22) $\text{cf}^*(x, \text{in}'(\text{y}), I)$ if and only if the interval $\{ t \in I : \text{loc}'(x)(t) \in \text{in}'(y)(t) \}$ $(= J)$ properly ends I

The result of this is a division of I into two *phases*, an initial phase in which the location of x is not interior to y and a final phase J in which the location of x is

Table 11.2. Five modes in Kracht (2002)

Mode	Meaning	Common case
static	'at'	locative, essive
coinitial	'from'	ablative, elative
cofinal	'to'	allative
transitory	'through'	perlative
approximative	'towards'	versative

interior to *y*. This represents English *into* as well as the Finnish illative. The complex preposition *out of* and the elative represent the opposite situation, where *J* properly begins the interval *I*.

Fong (1997) has also stressed this *diphasic* character of cases like the illative and elative. But while Kracht basically treats the transition as spatiotemporal, Fong makes the point that we need a more abstract notion of phases, on the basis of examples like the following, in which there is no literal movement:

(23) a. *silta San Francisco-on* (Fong 1997: 2)
bridge San Francisco-ILL
'a/the bridge into San Francisco'

b. *Toini rupeaa luke-ma-an* (Fong 1997: 51)
Toini begins read-INF-ILL
'Toini begins ("into") reading'

c. *Tuovi unohti kirjan auto-on* (Kracht 1997: 2)
Tuovi forgot book car-ILL
'Tuovi left the book in ("into") the car'

Fong argues that the illative case (and English *into* to a certain extent) expresses an opposition of two phases over an interval. The interval does not have to be temporal: it can also be the length of a bridge, as in (23a), with a phase outside the city followed by a phase inside. The transition can also be a non-spatial transition, as in (23b), with a non-reading phase preceding a reading phase. (23c) shows how the illative can even encode the opposition between 'not left in the car' versus 'left in the car'.

Kracht (2003) explains cases like (23c) by distinguishing semantic from syntactic selection and leaving the mode or the configuration semantically empty:

(24) a. *Hän menee laivalta.*
He walks ship-ABL
'He is going/walking from the ship.'

b. *Hän löysi rahansa laivalta.*
He found his money ship-ABL
'He found his money on ("off") the ship.'

978–0–19–920647–6 12-Spencer-ch11 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 181 of 182 May 16, 2008 18:58

c. *Tämä näyttää laivalta.*
This resemble ship-ABL
'This looks like/resembles ("off") a ship.'

In (24a) both the coinitial mode of the ablative and its 'on' configuration are semantically meaningful. In (24b), however, the verb selects a location. The coinitial mode only has a syntactic function, and is semantically empty. In (24c), both mode and configuration have null semantics, because the verb semantically selects for a thing. Kracht argues that disentangling syntax and semantics makes it easier to explain the differences that we find between languages in the use of modes and configurations.

11.4 Conclusion

We have only been able to give a sketch of how formal tools are being used in the study of the meaning of case. There are many semantic aspects of case that lend themselves to such a treatment, such as argument structure, quantification, aspect, and space, allowing us to gain a more precise and systematic understanding of the way case works. We can only hope for a more intense interaction between the empirical and formal approaches to this domain.

Acknowledgement

We gratefully acknowledge the financial support of the Netherlands Organisation for Scientific Research (NWO) for the research reported here (grant no. 220-70-003 'Case cross-linguistically' for both authors, grants no. DN 30-609 'Incremental interpretation of case and prominence' and 360-70-222 'Animacy' for the first mentioned author only).

PART II

MORPHOLOGY OF CASE

CHAPTER 12

CASE AS A MORPHOLOGICAL PHENOMENON

ANDREW SPENCER

12.1 Preliminaries: case in syntax and morphology

In most modern approaches to grammar the term 'case' is systematically ambiguous between '(inflected) form of a nominal word' and 'property of a noun phrase (determiner phrase)'. In a typical, well-behaved system we find that the lexical head noun of a phrase is, say, 'in the accusative' and this means that the phrase itself bears accusative case. However, as many of the articles in this handbook testify, this simple relationship between morphological case ('m-case') and syntactic case ('s-case') is often violated. In this chapter I will be concerned mainly with the morphological expression of case. However, it will be necessary to bear in mind the m-case/s-case distinction at various points.

12.2 Types of morphological case system

Cases are usually suffixes. The *World Atlas of Linguistic Structures* (WALS) survey (Haspelmath et al. 2005: 210–11) records a number of instances in which

cases are realized by prefixes but suffixation is by far the commonest type of realization. Chukchee (Muravyova 1998) has two comitative cases, both marked by a circumfix (*ge-...-te* and *ga-...-ma*). Gaelic marks genitive case by consonant mutation (Fife and King 1998: 489). *WALS* lists four languages which express case relations with tone alternations and two Nilotic languages (Dinka, Nuer) in which case is expressed by stem alternations (see also Blevins, this volume, Chapter 13). Vocative case is particularly prone to unusual marking, and often differs completely from other cases (see Daniel and Spencer, this volume, Chapter 43, for examples). One particularly difficult question concerns case marking by means of clitics (*WALS*: 211). In a great many languages formatives which have case functions attach to the edge or to the first constituent of the NP and are thus best thought of as clitic adpositions. However, a postposition in a head-final language tends to look like a case suffix and is often so described.

The one language type whose grammars uncontroversially have to appeal to a case feature is the traditional inflecting type, as represented by Latin. Latin has five cases but several distinct inflectional classes (traditionally five). Case markers cumulate with inflection class, that is, a single inflectional ending simultaneously conveys the case and the declension class. As a result no single morph ever uniquely expresses a case. We therefore need the m-case feature to express the fact that *mensae* 'table(GEN SG)' and *equī* 'horse(GEN SG)' express exactly the same grammatical functions. Independently, Latin also needs a syntactic [Case] attribute, s-case, for instance, to describe case agreement within NPs. We can contrast Latin with Japanese in this respect (Spencer and Otoguro 2005: 134f.). Japanese expresses grammatical relations by means of particles such as *ga* 'NOM', *o* 'ACC', *no* 'GEN', *ni* 'DAT'. Morphologically the particles are tightly bound clitic postpositions, not affixes. They can be replaced by emphatic particles such as *sae* 'even', *dake* 'only', or the topic marker *wa*. In certain frameworks it might be necessary for theory-internal reasons to describe Japanese syntax by appeal to some s-case attribute. However, there is no m-case in Japanese.

Beard (1995) has argued that a [Case] attribute is sanctioned in a formal description if and only if that attribute is needed to generalize across cases independently of form, as in Latin. We can generalize this to syntax, if an s-case feature is required to capture instances of case agreement and perhaps certain types of government phenomena. Spencer and Otoguro (2005) call this generalization 'Beard's Criterion'. What it means is that there is no need for a [Case] feature in a grammar except in order to generalize across distinct forms which have the same function. By Beard's Criterion a [Case] feature for Japanese, for instance, would be completely superfluous since the case particle (or clitic postposition) can be identified uniquely in terms of its form. These issues are discussed in detail in Otoguro (2006).

978–0–19–920647–6 13-Spencer-ch12 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 187 of 199 May 16, 2008 18:58

A language can have m-case without needing s-case, for instance if all syntactic dependencies can be stated purely in terms of the m-case forms. It's possible in principle for a language to have s-case without m-case, for instance, where case relations are marked by means of invariant adpositions which have to be repeated on modifiers, but I'm not aware of any clear examples of this.[1]

A grammatical description requires the m-case attribute whenever case is cumulated with some other nominal category such as declension class, number, definiteness, possessor agreement. In Latin case cumulates not only with inflection class but also with number, e.g. *equ-ī* 'horse.genitive.singular' vs. *equ-ōrum* 'horse.genitive.plural'. Yimas (Sepic-Ramu, New Guinean) has one case marker, with two allomorphs *-n*/*-nan*. Broadly speaking, *-n* is found with singular forms and *-nan* is found with non-singulars (duals, plurals), though there are various complications depending on noun class and the phonology of the plural form itself (Foley 1991: 166). Thus, the grammar of Yimas needs the m-case feature to account for the distribution of its only marked case. In Moksha Mordvin (Uralic, Feoktistov 1966: 204) the case form depends on definiteness/number (in the indefinite form it is only the nominative that has a separate plural form):

(1) Noun inflection in Moksha Mordvin: *alaša* 'horse'

	indefinite		definite	
	SG	PL	SG	PL
NOM	alaša	alaša-t	alaša-s^{j}	alaša-tne
GEN	alaša-n^{j}		alaša-t^{j}	alaša-tne-n^{j}
DAT	alaša-ndi		alaša-ti	alaša-tne-ndi

In Udmurt (Uralic), some cases cumulate plural number and others cumulate possessor agreement (Csúcs 1998: 282f.). The data in (2) show that some cases assume different forms with possessed forms of nouns (possessor agreement is marked with a suffix):

(2) Udmurt cases with possessed nouns

	1sgPx	2sgPx	3sgPx
Nominative	-(j)e, -ï	-(j)ed, -ïd	-(j)ez, -ïz
Genitive	-elen/ïlen	-edlen/ïdlen	-ezlen/ïzlen
Accusative	-me	-te, -de	-se, -ze
Inessive/Illative	-am	-ad	-az
Elative	-ïš^{j}tïm	-ïš^{j}tïd	-ïš^{j}tïz

[1] In Principles and Parameters models all DPs bear a value for the 'Abstract Case' feature even where there is no case morphology. However, this special Case feature is related only indirectly to standard interpretations of case and often fulfills the function of a proxy for reference to grammatical relations. See Bobaljik and Wurmbrand (Chapter 3).

Note that the inessive/illative case distinction is neutralized in possessed nouns, an instance of a change in paradigm structure conditioned by a morphological property (possessor agreement).

12.3 Case morphology and grammatical theory

An important theoretical aspect of case morphology is the phenomenon of case syncretism, where by a single word form occupies two or more cells in the paradigm (see Baerman, this volume, Chapter 14). Stump (1993, 2001a) discusses case syncretisms in Sanskrit, arguing that some, at least, support realizational models over morpheme-based models (specifically his Paradigm Function Morphology). In a realizational model, morphological rules define the word forms which occupy the cells of a lexeme's inflectional paradigm. The affixes themselves are not lexical entries with a fixed feature content. Stump argues that certain of the Sanskrit declensional syncretisms have to be handled by rules of referral, which stipulate that the form occupying a given cell is identical to some other word form. Sanskrit rules of referral can override and be overridden by rules of exponence.

Case morphology has played a role in syntactic debate. The Principles and Parameters model appealed to the Mirror Principle, according to which the linear order of affixes is defined by the order of syntactic processes. This principle was criticized by Spencer (1992) on the grounds that it failed to handle case marking. Within the noun phrase, possessor agreement is a relation which is entirely internal to the phrase, while case marking is imposed externally by a governing case marker. Therefore, by the Mirror Principle, possessor inflections should appear nearer the noun root than case markers, as is the case in Hungarian. However, in many languages, such as Finnish, we find the opposite order. However, in Distributed Morphology (Halle and Marantz 1993) this ceased to be a problem: case and agreement morphology were introduced as part of the mapping between syntactic structure and morphological structure, and so were allowed to violate the Mirror Principle.

The case stacking of Kayardild and other Australian languages has been analysed within Stump's Paradigm Function Morphology by Sadler and Nordlinger (2006), who propose a way of integrating this approach into Lexical-Functional Grammar (see Butt, this volume, Chapter 2 for other proposals within LFG).

Finally, interesting proposals for integrating morphological case marking with the specification of grammatical functions have been made within two very closely related frameworks, by Kiparsky (2001) and by Wunderlich (e.g. Wunderlich and Lakämper 2001). For a brief overview of these proposals see Butt (2006) and Primus (this volume, Chapter 17.).

12.4 Case-marked categories other than nouns

In languages such as Latin attributive modifiers (usually adjectives) agree in case with their head noun. In a number of languages a possessor noun which is formally in the genitive case will additionally take the case marker of its head, leading to the prototypical instances of *Suffixaufnahme* or double-case marking (as discussed in most of the papers in Plank 1995b). Sometimes the newly incorporated suffix supplants, rather than co-occurs with, the original suffix. See Chapter 15 and Plank (1995a: 20) for an example from Classical Armenian.

In many languages, including some which lack case agreement such as Hungarian, it's common to find that adjectives can be inflected as though they were nouns provided they function as nouns. Case marking is also very common on verbs, especially, but not exclusively, on non-finite or nominalized forms of verbs, generally giving rise to words functioning as adjuncts (often called 'gerunds' in the descriptive literature).

12.4.1 Case marking of verbs

The use of semantic case markers to form 'gerunds' is widespread, e.g. Chukotkan (Chukchee, Muravyova 1998: 534, Spencer 1991: 28–9), Alutor (Kibrik, Kodzasov, and Muravyova 2004: 265), Daghestan languages (e.g. Archi, Kibrik 1998), and Australian languages (see Dixon 1980: 381f. 458f. on the purposive gerund derived from a dative/purposive case marker and also Dench, this volume, Chapter 52 on Nyamal). In Quechua subordinate clauses functioning as the object of a matrix verb such as 'want', 'see (that)' are headed by non-finite (nominalized) verb forms marked with accusative case (Adelaar and Muysken 2004: 226):

(3) *rima-y-ta* *xala.yu-ru-n*
speak-INF-ACC begin-PRF-3SG
'He began to speak'

12.4.2 Cases in Kayardild

Evans' (1995a) description of the Tangkic, non-Pama-Nyungan language Kayardild identifies four types of case. The first two, adnominal case (realizing grammatical relations within NPs) and relational case (realizing grammatical relations at the clause level) are the vanilla varieties of case. However, five of the relational cases, locative, ablative, proprietive, oblique, allative, can be used with the function of

indicating or agreeing with various tense/mood/aspect categories on the main verb. Evans calls this 'modal case'. Within certain types of subordinate clause, NPs may take the locative or oblique cases to signal when a subordinate clause is an argument of the main verb or as a kind of switch-reference. This is called 'complementizing case'. Finally, Kayardild has another set of cases derived from verbs, called 'verbal case'.

12.4.2.1 *Modal and complementizing case*

Modal case is illustrated in (4) (Evans 1995a: 107–8):

(4) a. *ngada* *warra-ja* *ngarn-kir*
I.SG.NOM go-ACT beach-ALL
'I am going/have gone to the beach'
b. *ngada* *warra-ju* *ngarn-kiring-ku*
I.SG.NOM go-POT beach-ALL-MPROP
'I will go to the beach'
c. *ngada* *warra-jarra* *ngarn-kiring-kina*
I.SG.NOM go-PAST beach-ALL-MABL
'I went to the beach'
d. *ngada* *warra-da* *ngarn-kiring-inj*
I.SG.NOM go-POT beach-ALL-MOBL
'I would like to go to the beach'

Oversimplifying somewhat, modal case is attached to any non-subject NP to help express the TMA properties expressed on the verb (i.e. as co-realizations of the categories 'potential', 'past', 'desiderative'). The 'actual' tense form in (4a) does not require overt modal case marking.

Complementizing case is shown in (5):

(5) *ngada* *murnmurdawa-th*
I.SG.NOM rejoice-ACT
nginjin-inja *thabuju-ntha* *thaa-thuu-nth*
my-COBL elder.brother-COBL return-POT-COBL
'I am glad that my big brother is coming back'

In (5) each member of the subordinate clause is marked by the oblique case to signal the switch in subject. The oblique case must be selected here because the subject of the subordinate clause is third person.

12.4.2.2 *Verbal case*

Verbal case is a typologically unusual (if not unique) type of case found in Kayardild and the closely related Lardil. Verbal cases are mainly derived historically from

978–0–19–920647–6 13-Spencer-ch12 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 191 of 199 May 16, 2008 18:58

incorporated verbs and these retain some of their morphological verbal properties while still functioning as cases (Evans 1995a: 163–84): they attach exclusively to nouns (including nominalized forms of verbs) and produce noun forms, but the verbal case endings themselves behave morphologically like verbs. For instance, they belong to one of the two conjugation classes (indicated by theme elements *-tha*/*-ja*) and can take the middle voice marker *-a*. In addition, they agree with the main verb in tense, mood, and polarity:

(6) *ngada* *warra-ju* *dathin-kiiwa-thu* *ngilirr-iiwa-thu*
I.SG.NOM go-POT that-V.ALL-POT cave-V.ALL-POT
'I will go to that cave'

Before taking regular nominal case suffixes, nouns marked by verbal case have to be 're-nominalized', with one of the two nominalization suffixes, as seen in the example of verbal denizen case in (7) provided by Evans and Nordlinger (2004):

(7) *ngijin-mirdi-n-da* *dul-wirdi-n-da* *jardi-y*
my-V.DEN-NMZR-NOM place-V.DEN-NMZR-NOM mob-NOM
'the people staying at my place'

Note, too, that in this example the verbal case even has the adnominal case function, which is not even available to all types of relational case.

However, the NPs thus marked remain NPs. As seen in (6)–(7), the verbal cases have to be repeated on modifiers within the NP, just like 'normal' cases. Moreover, verbal case-marked NPs have all the distributional properties of NPs and none of the distributional properties of verbs. The verbal cases thus induce a particularly clear instance of a mismatch between morphological category (verb) and syntactic category (noun).

12.5 Syntagmatic aspects of case marking

Aspects of the linear ordering of cases are discussed in Chapter 15, so here I simply summarize the salient phenomena. Many languages show 'compound cases', that is local (spatial) cases formed from two elements, one of which identifies an orientation and the other of which specifies a direction or position. Thus, compound cases translate notions such as 'from inside', 'onto the surface of', 'at the front of'. Highly elaborated spatial case systems built in this manner are found in the languages of Daghestan (see Daniel and Ganenkov, this volume, Chapter 46; see also Comrie and Polinsky 1998, Kibrik 1998).

The order of case markers with respect to other types of inflection is sometimes variable (see Malchukov, this volume, Chapter 44 for several different types of example). A particularly intriguing form of variability is free variation, found in systems which have fairly recently grammaticalized properties such as plural number and case (see Moravcsik, Chapter 15, on Komi Zyryan). Luutonen (1997) discusses particularly interesting examples in the Uralic language Mari, of which there are two principal varieties, Meadow Mari and Hill Mari. Depending on dialect the plural is marked by *βlä* (Hill Mari) or by either *βlak* or *saməč*. Nouns inflect for possessors and for one of nine cases. The principal patterns of variation given by Luutonen (1997: 34) are shown in (8) ('Pl' = plural, 'Px' = possessor, 'C' = case).

(8) a. stem-Pl-Px ∼ stem-Px-Pl

b. (i) Genitive stem-Px-C (stem-C-Px)
Accusative stem-Px-C (stem-C-Px)

(ii) Inessive stem-Px (stem-Px-C)
Illative stem-C-Px (stem-Px-C)
Lative stem-C-Px (stem-Px-C)

(iii) Dative stem-C-Px ∼ stem-Px-C
Modal stem-C-Px ∼ stem-Px-C

c. (i) Genitive stem-Px-Pl-C ∼ stem-Pl-Px-C (stem-Pl-C-Px)
Accusative stem-Px-Pl-C ∼ stem-Pl-Px-C (stem-Px-C-Pl)

(ii) Inessive stem-Px-Pl-C ∼ stem-Pl-C-Px (stem-Pl-Px-C)
Illative stem-Px-Pl-C ∼ stem-Pl-C-Px (stem-Pl-Px-C)
Lative stem-Px-Pl-C ∼ stem-Pl-C-Px

(iii) Dative stem-Px-Pl-C ∼ stem-Pl-Px-C ∼ stem-Pl-C-Px

In these tables an alternate order in parentheses is possible but less preferred, otherwise we have apparently completely free variation. To a large extent variability depends on the case marker. The greatest variability is found with the dative. The one set of orders which is not possible is one where the case suffix precedes the plural suffix. Luutonen (1997: 36) explains this in terms of the grammaticalization path of these constructions. Otherwise, all logically possible orders are found. Some examples are given in (9), (10):

(9) Mari *olma* 'apple' accusative, lative cases

a. *olma-m-βlak-əm*
apple-1SG.PX-PL-ACC
olma-βlak-em-əm
apple-PL-1SG.PX-ACC 'my apples'
stem-Px-Pl-C ∼ stem-Pl-Px-C

b. *olma-m-βlak-eš*
apple-1SG.PX-PL-LAT
olma-βlak-eš-em

apple-PL-LAT-1SG.PX 'in my apples'
stem-Px-Pl-C ~ stem-Pl-C-Px

(10) Mari *joltaš* 'friend', dative case
joltaš-em-βlak-lan
friend-1SG.PX-PL-DAT
joltaš-βlak-em-lan
friend-PL-1SG.PX-DAT
joltaš-βlak-lan-em
friend-PL-DAT-1SG.PX 'to my friends'
stem-Px-Pl-C ~ stem-Pl-Px-C ~ stem-Pl-C-Px

In other Uralic languages the linear position of the case markers is fixed but depends on the case. Thus, in Mordvin the order of formatives is stem-Pl-Px-C in the nominative, genitive, dative forms, but stem-C-Px in the ablative, inessive, elative, and illative forms (these cases don't distinguish singular from plural). There is no free variation (Feoktistov 1966).

In the Kayardild examples discussed above we saw that a single nominal may receive more than one case marker, by virtue of the different functions that case markers are put to, so-called 'case stacking' (see Malchukov, Chapter 44). Case-stacking is particularly prevalent in Australian languages (see Malchukov, Chapter 44, Dench, Chapter 52, and Dench and Evans 1988, for an Australian survey), but it is found elsewhere too, e.g. Korean.

12.6 Problematic phenomena

The morphology of case raises a variety of interesting conceptual and analytical problems. I summarize the most salient of these. One prevalent problem is the distinction between 'zero' marking and having no marking at all, discussed in Malchukov and de Swart, Chapter 22, Amberber, Chapter 51 (see also Comrie 1986). Here I mention a number of problems that tend to get neglected in the literature.

12.6.1 Case inventories: 'core' vs. 'non-core' grammatical functions

Although it is sometimes thought that the principal rationale behind case marking is the identification of core arguments (subject, objects), there are many languages

978–0–19–920647–6 13-Spencer-ch12 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 194 of 199 May 16, 2008 18:58

in which cases are solely used to mark non-core grammatical relations, such as locations (see König, Chapter 50). Yimas (Foley 1991), for instance, has one oblique case used for a variety of purposes, but not for marking core arguments. A number of Andean languages have a wider selection of non-core cases. The Chibchan language Muisca has three main spatial case markers but normally subjects and objects remain unmarked (Adelaar and Muysken 2004: 9).

Generally, languages with purely spatial or oblique case marking don't fit Beard's Criterion. Where the cases are clearly affixal they thus behave more as though they were fused postpositions rather than canonical cases, though if a marker has the properties of an affix then it's generally called a case rather than a 'postposition' or 'particle'. This descriptive indeterminacy can lead to difficulties of analysis. For instance, the Bantu languages lack a case system, yet in Swahili (and other languages) we find a special locative form in *-nī nyumba-ni* 'to, at, by, in, from the house', *meza-ni* 'to, at, by, on, from the table', *kikapu-ni* 'to, at, in, out of, from the basket', and so on. This is an affix (for instance, it is in the stress domain of the word, Ashton 1944: 18). In a description which doesn't adhere to Beard's Criterion, it isn't clear why such a suffix doesn't constitute a case form.

12.6.2 Case inventories and outliers

In this subsection we look at defective systems in which only very restricted subclasses inflect for case. Case systems often decline with historical change, leaving vestigial forms (Arkadiev, Chapter 47). In (11) below we see the words 'boy', 'girl' in Hindi inflected for three vestigial cases (see also Malchukov, Chapter 44 on 'layered cases'):

(11) Inflected noun forms

	Singular	Plural	Singular	Plural
Direct	laRkaa	laRke	laRkii	laRkiyãã
Oblique	laRke	laRkõ	laRkii	laRkiyõ
Vocative	laRke	laRko	laRkii	laRkiyo
	'boy' (masculine)		'girl' (feminine)	

Most adjectives in Hindi are indeclinable, but a subset agree with nouns for case, number, gender. Thus, we have from the adjective *acchaa* 'good', *acche laRke* 'good boys.DIR/good boy.OBL', *acchii laRkiyãã* 'good girls.DIR', *acche laRkõ* 'good boys.OBL', *acchiyõ laRkiyõ* 'good girls.OBL'. Hindi is often said to have another case system, based on a set of clitic postpositions which express grammatical function roles of subject and object as well as various adverbial functions. However, these clitic formatives have none of the other properties of canonical case and completely fail to meet Beard's Criterion (see Spencer 2005 for detailed discussion).

Another type of attrition of the Indo-European case system is found in German. Determiners and phrase-initial adjectives are inflected for number and nominative,

accusative, genitive, dative. Nouns show number but not case distinctions. The only reliably case-marked form is the dative plural, which ends in *-n* (unless the plural form already ends in *-n*): *Mann* 'man', *Männer*–plural, *Männern* – dative plural. Masculine/neuter nouns (and some feminine names and kin terms such as *Mutti* 'Mummy') have a genitive singular form in *-(e)s*, though its functions are often taken over by the preposition *von* 'of'. What this means is that case is marked principally on the NP-initial determiner. Case marking on nouns is vestigial, and it certainly can't be said that nouns distinguish four cases. Instead, a noun's paradigm has to be described as {genitive singular, dative plural, other}.

Bulgarian has lost its case system except for a residual distinction between subject and direct object in the personal pronouns (Scatton 1984: 372f). However, some nouns and masculine gender adjectives have a productive vocative form (see Daniel and Spencer, Chapter 43). In addition, the written language distinguishes (is supposed to distinguish) subject and object forms of the masculine singular definite article (the definite article takes the form of a clitic attached to the first word of a NP): for subjects *-ət*, for all other functions *-ə*. The prescription is more honoured in the breach than the observance (Scatton 1984: 316; Gergana Popova, pers. comm.), but nevertheless it illustrates the way that somewhat arbitrary distinctions can remain after the breakdown of a case system. (See also Arkadjev, Chapter 47.)

12.6.3 'Case' marking on pronominal clitics

It is common in languages that have lost a case system, such as English, the Romance languages, Bulgarian, Macedonian, to retain vestiges of case in the pronoun system, especially the clitic pronouns. This raises interesting problems of description. It does not seem reasonable to say that, for instance, Bulgarian or the Romance languages retain a case system by virtue of the distinctions of 'case' made in their pronominal clitic systems, especially for Macedonian where the pronominal 'clitics' are really verbal affixes. It is often said that English pronouns distinguish case. For reasons why this is the wrong analysis see Hudson (1995).

12.6.4 'Non-autonomous' ('virtual') cases

Occasionally, in otherwise agglutinating languages we find evidence that a case forms part of the paradigm of a noun but there is no dedicated case marker which uniquely identifies that case. For instance, one might argue that the syntax of Estonian appeals to a notion of 'accusative case' but no nominal in Estonian has a form that can be uniquely identified as 'accusative'. Chukchee is a clearly ergative language, but no nominal in Chukchee has a case form which is uniquely identifiable as ergative. Class II nouns (broadly speaking human proper nouns) use

a form which is identical to the locative, while Class I nouns (all others) use a form which is identical to the instrumental (Spencer 2006). Mel'čuk (1986) refers to such 'virtual' cases as 'non-autonomous cases'.

12.6.5 Segmentation problems

We frequently encounter difficulties in knowing exactly how to segment case morphs. We noted that Daghestan languages have 'compound cases'. Comrie and Polinsky (1998) contrast these languages with the type of case system found in Uralic languages such as Finnish or Hungarian. The local cases of Finnish are shown in (12). Comrie and Polinsky distinguish 'in' from 'on' orientation in this table, and I have added the traditional case names. (Theirs is not the traditional way of presenting these case forms; but see Abondolo 1998: 157 for a more thoroughgoing classification along these lines):

(12) Finnish local cases

	'in'	'on'
essive	-ssa (=inessive)	-lla (=adessive)
allative	-(h)Vn (=illative)	-lle (=allative)
ablative	-sta (=elative)	-lta (=ablative)

The 'V' of the suffix *-(h)Vn* indicates a vowel determined by the root, (though this suffix has the partially suppletive allomorphs determined by the number of syllables in the stem and the type of stem ending; see Karlsson 1982: 103–5 for details). Although there are vestiges of suffixes *-s-* 'in', *-l-* 'on' Comrie and Polinsky (1998: 107) remark 'it is not clear that the decompositional analysis at a synchronic level has any real advantage over saying there are just six unanalyzable case suffixes'. Of course, the real question here is 'what criteria should morphological theory deploy to resolve such a question?'

A further problem of segmentation arises from partial grammaticalization of postpositions attached to inflected noun forms, typically denominal postpositions attaching to genitive case forms. A good example of the indeterminacy this can cause is provided by Shiffman's (1983) description of the Dravidian language Kannada. The genitive is mainly formed by adding *-a* to the base form though the genitive suffix is often lost through morphophonemic changes. Case endings (dative, instrumental/ablative, locative, vocative) are added either to the base form or to the oblique stem (depending on factors such as gender). The oblique stem is often, but not invariably, identical to the genitive case. Shiffman (1983: 182 fn.4) raises the question whether the genitive can be treated as a stem form, arguing that this is not possible because some nouns have a distinct genitive. But it is problematical to treat the oblique stem itself as the genitive form, because then a number of cases would involve double case marking (i.e. Suffixaufnahme). In fact

978–0–19–920647–6 13-Spencer-ch12 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 197 of 199 May 16, 2008 18:58

what frequently seems to be happening in such languages is that cases develop from postpositions added to a genitive case, but that genitive then gets reanalysed as a stem form, without necessarily losing its original function as a genitive case form. In such grammars we therefore need to set up an oblique stem and a genitive case and a default rule or principle which defines the genitive as the unaffixed oblique stem (this is also needed for a number of Daghestan languages except that the oblique stem doubles as the ergative case; see Kibrik 1998 on Archi). The distinct genitive forms then override this default. (See §12.6.8.2 below on Estonian.)

12.6.6 Inflection or derivation?

The 'structural' cases (nominative, accusative, ergative, absolutive, possibly also genitive, dative) are almost always taken to be inflections, but 'semantic' cases often have the character of derivational formatives, in that they add a clearly defined meaning to the noun, much in the way that a preposition adds meaning to a noun phrase. This can make it difficult to decide whether we are dealing with a case-inflected form of a lexeme or a new lexeme with adverbial function. The problem is particularly acute in cases of on-going grammaticalization. Evans (1995a: 180f.) discusses this question with respect to verbal case in the Tangkic language, Kayardild, showing that this particular kind of case is almost certainly inflectional in that language, though in the closely related language Yukulta it's still essentially derivational. Another useful discussion of this problem, this time for Hungarian, is provided by Kiefer (1987). See also Malchukov's discussion of 'distributed case' in Koasati (Chapter 44).

Another difficult area can arise when we have Suffixaufnahme, particularly when a genitive-marked possessor agrees in case with its head. When the phenomenon was first observed (for Old Georgian) by Bopp it was argued that the genitive must actually be a kind of adjective, so that the genitive inflection was really a derivational adjectivizer (Plank 1995a: 4).

12.6.7 Case endings as nouns in Hungarian

Hungarian nouns take between seventeen and twenty-eight case suffixes (depending on the analysis), but Hungarian pronouns lack case-marked forms. Instead, a default form of the case affix is taken and inflected as though it were a possessed noun. Thus, the dative case suffix *-nek*/*-nak* gives *Ferenc-nek* 'to Ferenc', *László-nak* 'to László', but *nek-em*, *nek-ed*, *nek-e* . . . 'to me, you, him/her' (cf. *gyerek-em*, *gyerek-ed*, *gyerek-e* 'my child, thy child, his/her child'). Phenomena such as these raise interesting questions about what exactly a 'case suffix' is in Hungarian (Spencer, forthcoming).

12.6.8 Case and morphosyntax

12.6.8.1 *Suspended affixation: Turkic and other Eurasian languages*

Lewis (1967: 35) describes a phenomenon in Turkish grammar which he calls 'suspended affixation', in which a single affix is attached to the final member of a coordinated phrase and taking scope over the whole phrase. In (13)–(14) we see examples involving case:

(13) a. *sihhat-ta afiyet-te*
health-LOC well-being-LOC
b. *sihhat ve afiyet-te*
health and well-being-LOC
'in health and well-being' (N.B. vowel harmony)

(14) *tebrik ve teşekkür-ler-im-i*
congratulation and thank-PL-1SG.PX-ACC
(cf. *tebrik-ler-im-i* congratulation-PL-1SG.PX-ACC 'my congratulations')

The property of taking wide scope is generally associated with clitics, not affixes. Yet it is generally assumed in Turkish grammar that the cases are part of the inflectional paradigm of a noun lexeme.

12.6.8.2 *Cases and agreement: Estonian's 'four last cases'*

An interesting mismatch is seen with the 'four last cases' of Estonian. Estonian adjectives agree with their head nouns in case and number:

(15) Paradigm for *noor inimeme* 'young person'
(adapted from Uuspõld and Valmet 2001: 192)

	SG		PL	
nominative	noor	inimeme	noore-d	inimese-d
genitive	noore	inimese	noor-te	inimes-te
<eight other cases . . . >				
terminative	noore	inimese-ni	noor-te	inimes-te-ni
essive	noore	inimese-na	noor-te	inimes-te-na
abessive	noore	inimese-ta	noor-te	inimes-te-ta
comitative	noore	inimese-ga	noor-te	inimes-te-ga

The 'four last cases' are attached to a stem form identical to the genitive case, which itself is unequivocally part of the case paradigm and not some special extra-paradigmatic form. An adjective agrees with a noun in one of these cases as though it were in the genitive, 'the wrong case'. Similar constructions are found in Romani and other Indo-Aryan languages (Koptjevskaja-Tamm 2000; Arkadjev, Chapter 47).

978–0–19–920647–6 13-Spencer-ch12 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 199 of 199 May 16, 2008 18:58

12.7 Conclusions

The complexity of morphological case systems provides a rich testing ground for morphological theory. Open questions include 'how do we distinguish (morphological) case markers from adpositions?'; 'how does case morphology relate to the inflection/derivation distinction?'; 'how do we integrate outlier cases or fossilized cases into a regular case system?'; 'how does morphological case marking relate to syntactic case specification?'. A particularly difficult set of problems is posed by application of Beard's Criterion, since this raises important questions about what we mean by 'case' in the first place.

CHAPTER 13

CASE AND DECLENSIONAL PARADIGMS

JAMES P. BLEVINS

13.1 Introduction

Case plays a central role in declensional paradigms, as reflected in the way that declensions are traditionally described as 'case paradigms'. Case is often the declensional feature with the largest and most heterogeneous inventory of values, and much of the structure of declensional paradigms derives from patterns within the case system. Some patterns involve oppositions between morphosyntactic case features, or between the forms that realize these features. However, other types of patterns are more 'purely morphological' (Aronoff 1994) in the sense that they relate sets of forms that do not comprise any kind of natural morphosyntactic class. It is the prevalence and intricacy of these morphological patterns that make case paradigms particularly relevant for general theories of morphology.

Three types of patterns are especially salient in case systems. The first involves a split within case systems. Languages with large case inventories frequently subdivide case forms into groups: one core set of 'grammatical' cases and one or more sets of 'secondary' or 'semantic' cases. The grammatical cases characteristically mark governed dependants such as subjects, objects, and indirect objects, as well as adnominal dependants such as possessors. The semantic cases tend to represent

spatial relations and other specific semantic properties. In many languages, the division between grammatical and semantic cases has a morphological reflex. Unmarked and fusional forms are far more likely to realize grammatical cases, whereas semantic cases are typically agglutinative, and are frequently based on a stem that corresponds to a grammatical case form. The grammatical/semantic case split tends to have a transparent historical basis, and semantic cases can often be shown to have arisen through the grammaticalization of postpositions or other formerly independent items, as Grünthal (2003) suggests for Finnic. Case systems exhibit other divisions, such as the split between ergative/absolutive and nominative/accusative patterns discussed by Silverstein (1976). However, the locus of this type of split lies outside the declensional system, as the paradigm of any given item will tend to follow a consistent nominative–accusative or ergative–absolutive pattern.

The second declension-internal pattern, which often cuts across other divisions, reflects the organization of case forms within a paradigm into form classes or 'cohorts'. Nominal cohorts, like the members of many verbal series, are often characterized by common affixal exponents and/or by stem alternations which are orthogonal to natural morphosyntactic or morphosemantic classes. In some languages, declension-internal form classes are defined by patterns of whole-word syncretism between morphosyntactically distinct paradigm cells. In others, notably in the Finno-Ugric and Daghestanian languages, sets of morphosyntactically heterogeneous cases share a common base. In many of these systems, the sets of case forms that share a common form element cannot be assigned 'general meanings' (à la Jakobson 1936) that make any substantive contribution to the specific grammatical meanings of individual forms. These patterns are, in short, intrinsically morphological, not morphosyntactic, even allowing a more abstract level of description, and to the extent that there are explanations for the patterns, the explanations tend to be historical not synchronic.

The third pattern concerns the implicational structure that arises from the 'relative informativeness' of the members of a paradigm. In languages with a rich inflection class system, the members of a paradigm are linked by multiple implicational relations. However, it is often the case that certain forms are more informative than others and play a 'diagnostic' role, marking the declension class of a nominal and implying other members of the paradigm. Diagnostic forms fall most frequently within the grammatical case system, again mainly for historical reasons. Older case forms tend to be less uniform morphologically, and are thus more likely to correlate uniquely with individual paradigms or inflection classes. In some systems, higher-level generalizations may also be of diagnostic value, with class-specific patterns of syncretism or contrast serving to identify the paradigm of a nominal.

These patterns are illustrated in section 13.2, which surveys a variety of case paradigms and clarifies some of the descriptive and theoretical issues that they raise.

978–0–19–920647–6 14-Spencer-ch13 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 202 of 218 May 16, 2008 18:58

Chapter 13 concludes with a discussion of implications of these patterns for models of morphological analysis.

13.2 The structure of case paradigms

Declensional paradigms associate two structured domains. The first, morphosyntactic, domain is defined by the distinctive features of a language. The morphosyntactic structure of an individual paradigm is determined by the feature bundles that characterize distinct 'cells' in that paradigm. This structure is conventionally represented in traditional paradigm tables, in which natural classes correspond to sets of cells with common features. The second, morphological, domain contains the word forms that occupy paradigm cells. There are two basic perspectives on the morphological structure of paradigms. Traditional 'word and paradigm' models treat paradigms as networks of word forms. On this view, the morphological structure of a paradigm reflects formal relations between members of the paradigm. The notion of 'exponence' subsumes those relations that correlate with natural classes of features. But traditional approaches also recognize purely formal relations that are orthogonal to feature classes. Patterns of this type are sometimes termed 'parasitic' (Matthews 1972) or 'morphomic' (Aronoff 1994). Within the post-Bloomfieldian tradition, the primary locus of part–whole relations is shifted from paradigms onto words, so that morphological structure arises in the analysis of individual words into component morphs. Morphological models developed within this tradition – as well as those influenced by it – accord paradigms no real status within the grammar, but treat them instead as collections of independently-defined forms that share common stems or patterns of exponence. The view that paradigms are derivative sets of forms is clear, for example, in Anderson's (1992: 134) definition of 'an item's paradigm' as 'the complete set of surface word forms that can be projected from its stem set by means of the inflectional Word Formation Rules of the language'.

The patterns exhibited by case paradigms in many languages bear in a fairly direct way on the evaluation of these alternative conceptions of morphological structure. In a sufficiently simple declensional system, it is often possible to regard morphological structure as derivative of morphosyntactic feature contrasts. However, languages with rich case inventories exhibit patterns and dependencies that appear to cut across classes of features and often contain 'units of form' at various levels that defy morphosyntactic analysis. Different types of patterns tend to arise in different types of case systems. As a rule, case systems with a small number of distinct forms exhibit more whole-word syncretism than stem syncretism,

Table 13.1. Case paradigm of 'weak' adjectives in German

	Masculine	Feminine	Neuter	Plural
Nominative	alte	alte	alte	alten
Accusative	alten	alte	alte	alten
Dative	alten	alten	alten	alten
Genitive	alten	alten	alten	alten

whereas the converse is true of systems with large case inventories. These broad tendencies can be illustrated by considering a number of case systems of ascending complexity.

13.2.1 Case syncretism

The smallest morphological case pattern involves a single opposition between a pair of distinct forms. This pattern is found in German, where 'weak' adjectives exhibit a binary opposition between a form in *-e* and a form in *-en*, as illustrated in Table 13.1. The paradigm of weak masculine nouns follows the same pattern as the masculine adjectives in Table 13.1. For example, the nominative singular form *Mensch* 'person' contrasts with a second form, *Menschen*, which realizes all of the remaining case–number combinations.

Adjective paradigms in the Nakh languages (Chechen, Ingush, and Tsova-Tush/Batsbi) also contain just a pair of distinct forms. As shown in the paradigm of *jogga* 'big' in Table 13.2, the nominative singular is realized by the stem form and the remaining case–number combinations are realized by a second, 'oblique' form, which in Ingush is marked by *-ča*.

Table 13.2. Basic noun and adjective paradigms in Ingush (Nichols 1994a: 96–99)

	Singular	Plural	Sing/Plur
Nominative	kuotam	kuotamaž	jogga
Genitive	kuotama	kuotamii	joggača
Dative	kuotamaa	kuotamažta	joggača
Ergative	kuotamuo	kuotamaž	joggača
Allative	kuotamaga	kuotamažka	joggača
Instrumental	kuotamaca	kuotamažca	joggača
Lative	kuotamağ	kuotameğ	joggača
Comparison	kuotamal	kuotamel	joggača
	'hen'		'big'

The descriptive challenges posed by the paradigms in Tables 13.1 and 13.2 arise at the level of morphosyntactic analysis. Without secure criteria for distinguishing feature neutralization from accidental homophony, it is not entirely clear how many cells to assign to the paradigm of *Mensch* or *jogga* (though see Chapter 14 on case syncretism in this volume for discussion). The form variation in these paradigms is, in contrast, relatively straightforward. In the paradigm of *jogga*, the form that realizes the nominative singular corresponds to the adjective stem, while the second form can be defined by suffixing the marker *-ča* to the stem. Hence, this paradigm can be defined by treating *joggača* either as the realization of an abstract 'oblique' case or else as a homophonous form that realizes multiple surface cases. In either event, there is no need to impose any further structure on the forms of *jogga*.

Evidence for paradigm structure comes from systems in which the formal patterns in a paradigm cross-cut feature classes in more intricate ways. In languages with a rich system of inflection classes, purely formal dependencies may even be of considerable diagnostic value. The paradigms in Table 13.3 illustrate three of the inflectional patterns for inanimate nouns in Polish. Of particular interest here are the syncretisms identified by the pairs of italicized forms. Each formal relation tends to hold between cells in a class of paradigms rather than between particular case exponents (though, as a reviewer notes, the identity of genitive singulars and nominative plurals in the neuter may be disrupted by stem alternations). The syncretism between locative and vocative singular in the paradigm of *świat* 'world' is also characteristic of nouns like *stół* 'table' with the syncretic form *stole*, as well as nouns like *brzeg* 'edge' with the syncretic form *brzegu*. The collapse of nominative plural and genitive singular in the paradigm of hard-stem neuters like *miasto* 'city' recurs in the paradigms of soft-stem nouns such as *zdjęcie* 'photo' and *narzędzie* 'tool', which have the syncretic forms *zdjęcia* and *narzędzia*. The identity of dative and locative singular forms in the paradigm of *szkoła* 'school' is characteristic of feminine declensions in general, as illustrated by the paradigms of *ulica* 'road',

Table 13.3. Inanimate noun paradigms in Polish (de Bray 1980: 263–80)

	Masculine *o*-stems		Neuter *o*-stems		Feminine *o*-stems	
	Singular	Plural	Singular	Plural	Singular	Plural
Nominative	świat	światy	miasto	*miasta*	szkoła	szkoły
Genitive	świata	światów	*miasta*	miast	szkoły	szkoł
Accusative	świat	światy	miasto	miasta	szkołę	szkoły
Dative	światu	światom	miastu	miastom	*szkole*	szkołom
Instrumental	światem	światami	miastem	miastami	szkołą	szkołami
Locative	*świecie*	światach	mieście	miastach	*szkole*	szkołach
Vocative	*świecie*	światy	miasto	miasta	szkoło	szkoły
	'world'		'city'		'school'	

978–0–19–920647–6 14-Spencer-ch13 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 205 of 218 May 16, 2008 18:58

armia 'army', *ziemia* 'earth' and *noc* 'night', which contain the syncretic forms *ulicy*, *ziemi*, *armii*, and *nocy*.

Class-specific syncretisms of the type exhibited in Table 13.3 cannot be subsumed under general patterns of case neutralization. Instead, these syncretisms are usually expressed by constraints that identify the realization of morphosyntactically distinct paradigm cells, whether in terms of the 'take-overs' of Carstairs (1984) or the 'rules of referral' of Zwicky (1985) and Stump (1993, 2001a). Take-overs and referrals both impose additional organization on a case paradigm – or on the rules that define the paradigm – by cross-referencing cells or rules.

13.2.2 Implication and reversal

The additional structure introduced by dependencies between cells in a paradigm can be captured in a variety of different ways. Take-overs and referrals are both 'directional' strategies in the sense that they define the form of one cell with reference to the form of another. This directionality often reflects a derivational perspective in which, as Stump (1993: 450) suggests, 'realization rules . . . encompass the individual steps by which an individual word is built up from the root of its paradigm'. The syncretisms in Table 13.3 can be understood equally well in terms of symmetrical 'paradigm structure constraints' (cf. Wurzel 1984) that identify pairs of cells as mutually informative, that is, as reliable predictors of the forms associated with each of the cells. This more static conception conforms to a traditional 'abstractive' perspective in which the members of a paradigm form a network of elements linked by implicational relations (Blevins 2006). From this standpoint, the syncretic cells in Table 13.3 are related by mutual implication – or, more formally, by a high 'mutual information' value, in the information-theoretic sense (Cover and Thomas 1991).

The mutual implication between pairs of identical forms represents a highly specific type of information about the elements of a case paradigm. An insight that underlies traditional word and paradigm approaches is that some members of a paradigm may be more informative than others, along at least two dimensions of 'informativeness'. Some aspects of a form identify the 'grammatical meaning' of the form, here its case and agreement features. Other aspects identify the inflection class or 'lexeme' that a form is affiliated with. A single form may mark both grammatical meaning and lexical class, as illustrated by *szkołę*, in which the ending *-ę* identifies an accusative form of an *o*-stem feminine noun. But forms may also identify the two different types of information in varying degrees. For example, an ending in *-ami* is diagnostic of instrumental plural in the paradigms in Table 13.3, but the instrumental plurals *światami*, *miastami*, and *szkołami* do not identify the class of the corresponding noun. In contrast, the grammatical meanings of *świat* or *szkoła* are not unambiguously signalled by any aspect of their form, given that a consonant-final form may realize the genitive plural, as in the case of *szkoł* or *miast*,

while an *a*-final form may realize the genitive singular, as in the case of *świata* and *miasta*. Pairs (or sets) of paradigm cells may likewise be asymmetrically informative. A nominative singular such as *świat* is diagnostic of a masculine *o*-stem noun, and implies a dative singular such as *światu*. However, a dative singular in *-u* may either be associated with a masculine noun, as in *światu*∼ *świat*, or with a neuter, as in *miastu*∼ *miasto*. These types of asymmetrical dependencies can again be modelled in either derivational or static terms, by means of derivational referral rules or take-overs, or in terms of information-theoretic notions such as the conditional entropy of a pair of cells.

For whole-word syncretism, there is perhaps not a great deal of difference between derivational and implicational strategies. The flexibility of an implicational perspective is brought out more clearly by other types of form dependencies in case paradigms. One striking pattern is exhibited by the 'external cases' that encode the syntactic function of nominals in Dinka. As Andersen (2002: 9) reports, the absolutive and oblique forms of many classes of Dinka nouns are distinguished by systematic patterns of 'tone reversal'.

> The oblique is distinguished from the absolutive in virtually all monosyllabic nouns that have a short vowel . . . and in most disyllabic nouns with the prefix à- and a short root vowel . . . The rule for such nouns is that if the absolutive has a low root tone . . . then the oblique gets a falling root tone, and if the absolutive has a high or falling root tone . . . then the oblique gets a low root tone.

Some examples that illustrate these alternations are given in Table 13.4. In each of the three patterns, it is the contrast between tones that is distinctive. There is no consistent association between tonal melodies and case values. Moreover, since these nouns are mostly simple monosyllables, it is unclear that the analysis of a case form involves any 'derivational steps' that a referral rule could access and invert in order to define the contrasting case form. In short, the operative relation

Table 13.4. Case alternations in Dinka (Andersen 2002: 9)

Alternation	Absolutive	Oblique	
low∼falling	Pìŋ	pîŋ	'spear'
	ɯ̀t	ɯ̂t	'house'
	àjìt	àjît	'chicken'
	àŋwèm	àŋwêm	'buttock'
high∼low	dít	dìt	'bird'
	báŋ	bàŋ	'chief'
	léc	lèc	'stick'
falling∼low	àŋâw	àŋàw	'cat'
	àrêw	àrèw	'tortoise'

978–0–19–920647–6 14-Spencer-ch13 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 207 of 218 May 16, 2008 18:58

between external cases in Dinka appears to hold between the surface forms that realize paradigm cells, and cannot be cast in terms of rules or devices that are taken to derive these forms. This type of dependency between forms fits well with a static conception of paradigm structure in which cells stand in implicational rather than derivational relations. For the nouns in Table 13.4, knowing the tonal melody of the absolutive allows one to predict the melody of the oblique. Knowing the oblique form either identifies the melody of the absolutive, or narrows the choice of melodies to one of two possibilities.

The tonal patterns in Dinka are reminiscent of the 'vowel reversal' that distinguishes indicative and subjunctive mood in the Romance paradigms discussed in Matthews (1991: 199). Estonian provides a more directly relevant parallel, as the 'theme vowels' in partitive singulars and 'stem' partitive plurals exhibit a similar pattern of vowel reversal. If a partitive singular ends in *-i*, as in the case of *`tooli* 'chair' in Table 13.10 below, then the stem plural ends in *-e*, as in *`toole*. Conversely, if the partitive singular ends in *-e*, as in *`lille* 'flower', the corresponding plural, *`lilli*, ends in *-i*. Like the tonal melodies in Dinka, the information value of the exponents *-i* and *-e* cannot be expressed by assigning them a definite case value, but only by recognizing their usefulness in predicting other forms of their case paradigm.

13.2.3 Stem syncretism

The Polish and Dinka paradigms in section 13.2.2 both involve dependencies between full word forms that occupy paradigm cells. Similar types of cross-paradigm dependencies may also hold between parts of word forms. In the general case, pairs or sets of forms share a common base, which differs from the basic root or stem of the paradigm. In a limiting case of this pattern, which is often termed 'Priscianic' or 'parasitic' syncretism, one word form constitutes the base for another form. Parasitic forms are often characteristic of complex case inventories that contain a set of grammatical case forms and a secondary set of semantic forms. The composition of parasitic forms also suggests the origin of some complex inventories, as the base tends to be drawn from the set of grammatical cases and the exponent of the semantic forms often corresponds to a grammaticalized adposition or other functional element.

A simple case of parasitic syncretism is found in Kabardian, where the oblique form of a noun provides the base for the instrumental form. This pattern is illustrated in Table 13.5.

In what Colarusso (1992) terms the 'unspecified' paradigm, there is no formal distinction between absolutive and oblique (or predicative) cases, so the instrumental marker could be treated as attaching to a basic stem such as *g^{y}aata* 'sword'. However, in the 'specified' paradigm, the instrumental marker *-k'^{y}a* is added to an oblique form marked by a final *-m*. Thus oblique *g^{y}aatam* provides the base for

Table 13.5. Absolute noun declensions in Kabardian (Colarusso 1992: §3)

	Specified Paradigm				Unspecified	
	Singular	Plural	Singular	Plural	Paradigm	
Absolutive	gʸaatar	gʸaatahar	šədər	šədhar	gʸaata	šəd
Oblique	gʸaatam	gʸaataham	šədəm	šədham	gʸaata	šəd
Instrumental	gʸaatamkʼʸa	gʸaatahamkʼʸa	šədəmkʼʸa	šədhamkʸa	gʸaatakʼʸa	šədkʼʸa
Predicative	gʸaataw	gʸaatahaw	šədəm	šədəm	gʸaata	šəd
	'sword'		'donkey'		'sword'	'donkey'

instrumental *gʸaatamkʼʸa*. The fact that the instrumental is a secondary formation based on a central set of cases leads Colarusso (1992: 51) to suggest that the instrumental is not a separate case, but rather a suffixed form of a case:

> Specified (definite or indefinite) nouns in Kabardian can take four cases: absolutive, oblique, instrumental, and predicative. Two of these, however, the instrumental, and the predicative, might be considered to be other than cases. The instrumental is actually a suffix on the oblique, while the predicative has a multitude of roles but can generally be seen as an adverb suffix or a complementiser, a sign that the noun comes from an underlying absolutive or oblique which has been lost due to syntactical complexities.

To some degree, the anomaly of the instrumental in Kabardian is due to the fact that it is a single, isolated form. In languages with larger case inventories, the status of secondary formations is often clearer. For example, the allative forms *kertaga* and *kuotamaga* (repeated from Table 13.2 above) underlie the secondary case forms of *kuorta* 'head' and *kuotam* 'hen' in Table 13.6. Since each secondary case has an invariant marker, the allative implies the form of the secondary cases, and each of the secondary cases identifies the allative base. Similar patterns are found in the other Nakh languages. Chechen contains six 'secondary cases' that are described as being 'derived from the allative' (Nichols 1994a: 24). Tsova-Tush likewise retains 'numerous complex case forms' of which 'the most frequent are locatives formed by adding the locative suffix -*ħ* to the allative or the allative II forms' (Holitsky 1994: 168).

The relation between the allative and the secondary cases in Ingush is, moreover, just one link within a larger network of implications. The allative singulars *kuotamaga* and *kertaga* are themselves based on the genitive singulars *kuotama* and *kerta*. These genitive singulars also provide a base (italicized in Table 13.6) for most of the other singular forms of *kuotam* and *kuorta*. The nominative plural *kuotamaž* is also based on the genitive singular, as in many noun classes. This pattern does not hold for all classes, as shown by the contrast between genitive singular *kerta* and nominative plural *kuortož*. However, in both paradigms, the nominative plural identifies the ergative plural and provides the base (set in bold in Table 13.6) for the

Table 13.6. Primary and secondary case forms in Ingush (Nichols 1994b: 95, n.d.)

	Singular	Plural	Singular	Plural
Primary Cases				
Nominative	kuotam	**kuotamaž**	kuorta	**kuortož**
Genitive	*kuotama*	kuotamii	*kerta*	kuortoi
Dative	*kuotama*a	**kuotamažta**	*kerta*a	**kuortožta**
Ergative	kuotamuo	**kuotamaž**	kertuo	**kuortož**
Allative	*kuotama*ga	**kuotamažka**	*kerta*ga	**kuortožka**
Instrumental	*kuotama*ca	**kuotamažca**	*kerta*ca	**kuortožca**
Lative	*kuotama*ğ	kuotameğ	*kertağ*	kuortuojeğ
Comparison	*kuotama*l	kuotamel	*kerta*l	kuortuojel
Secondary Cases				
Locative 2	kuotamagaħ	kuotamažkaħ	kertagaħ	kuortožkaħ
Ablative	kuotamagara	kuotamažkara	kertagara	kuortožkara
Ablative 2	kuotamagaħara	kuotamažkaħara	kertagaħara	kuortožkaħara
Translative	kuotamagağolla	kuotamažkağolla	kertagağolla	kuortožkağolla
	'hen'		'head'	

dative, allative, and instrumental plurals. These systematic form correspondences underlie the economy of an Ingush case paradigm, since a small number of forms identify the full case paradigm. It is this implicational structure that a traditional analysis exploits when it factors declensional classes into sets of exemplary paradigms and principal parts. In Ingush, the genitive singular and nominative plural are of obvious diagnostic value, though the forms based on the genitive singular or nominative plural are equally informative. Given the uniformity of case endings in Ingush, instrumental forms such as *kuotamaca* or *kertaca*, for example, uniquely identify their respective bases, *kuotama* and *kerta*. From a traditional perspective, then, nearly any form of an Ingush case paradigm provides a 'point of entry' into a network of implications that will identify the forms of other members of that paradigm.

The diagnostic value of nominatives and genitives in Ingush (as well as ergatives in many North Caucasian languages) is explicit in standard descriptions that classify these forms as 'stems'. A stem can, of course, be understood as a common 'element of form' that is shared by a set of interpredictable members of a paradigm, and in this sense stems are equally compatible with an abstractive or derivational view of paradigm structure. However, some accounts adopt or imply a narrower conception, in which stems are seen as basic units from which larger forms are constructed. In many instances, the postulation of stems serves mainly to avoid recognizing direct relations between surface word forms in a paradigm. The patterns in Dinka suggest that relations between surface forms cannot be avoided in general, and even in languages that exhibit pervasive stem syncretism, the introduction of abstract stems tends to obscure more than illuminate the structure

of a declensional paradigm. The question is not whether recurrent units of form are explicitly characterized as 'stems' within an analysis, since little hinges on this issue. The substantive point concerns the **status** of such units, however they are classified. On a constructive account, stems are building blocks from which larger forms are assembled, and which have a status independent of those larger forms. On an abstractive approach, stems are shared elements of form that are common to the elements of a case paradigm.

In the Ingush paradigms in Table 13.6, the genitive singular might be identified as an 'oblique' or 'second' stem, of the kind regularly posited for Daghestanian languages (Kibrik 1991, 1994a; Haspelmath 1993a). One could then construct an analysis on which the genitive singular is formally unmarked and thus realized by the oblique stem, while the dative, allative, instrumental, lative, and comparison cases are formed by suffixing a case ending to the oblique stem. Yet a constructive analysis along these lines creates a number of artefactual complications. The allative singular is marked by the exponent *-ga* and in this respect appears to pattern with the other singular case forms. But unlike the other forms, the allative provides a base for the secondary cases. If the secondary cases are taken to be based on the allative, the analysis admits relations between surface word forms and there is no obvious rationale for treating the genitive differently. But recasting the allative singular as a stem is artificial, as the former allative markers – which remain in complementary distribution with the other case markers – must be reclassified as stem formants, and the allative becomes another formally unmarked case, which is realized by the selection of a 'third' stem.

13.2.3.1 *Case series*

The difference between treating stems as a basis for form implications and treating stems as basic lexical units are clearer in languages with more intricate patterns of stem syncretism. Daghestanian languages exhibit one type of pattern, in which secondary cases are organized into a number of locative 'series'. The paradigms of the class I noun *wacc* 'brother' and the class II noun *kamyón* 'truck' in Table 13.7 display the structure of secondary series in Avar. Each series – including the 'grammatical' series – exhibits informative patterns of stem syncretism. The absolutive singular is the 'root' form of a noun, which underlies the absolutive plural as well as the singular and plural ergative forms. The root itself is of limited diagnostic value, as neither the ergative singular nor the absolutive or ergative plural can be predicted from the root in isolation. In contrast, the ergative singular and plural forms imply the form of the remaining grammatical cases. Although one cannot predict which formative will mark the ergative singular of a noun, given the ergative singular (either alone, or in conjunction with the ergative plural) one can segment the form into root and formative. The genitive and dative forms are in turn distinguished from the corresponding ergative by a final *-ul* or *-e*. The absolutive plural is again

Table 13.7. Noun declensions in Avar (Charachidzé 1981: §3)

	Singular	Plural	Singular	Plural
Absolutive	wacc	wáccal	kamyón	kamyónal
Ergative/Inst	wáccass	wáccaz	kamyónall`	kamyónaz
Genitive	wáccassul	wáccazul	kamyónall´ul	kamyónazul
Dative	wáccasse	wáccaze	kamyónall´e	kamyónaze
Superessive Series				
Locative	wáccassda	wáccazda	kamyónalda	kamyónazda
Allative	wáccassde	wáccazde	kamyónalde	kamyónazde
Ablative	wáccassdassa	wáccazdassa	kamyónaldasse	kamyónazdassa
Translative	wáccassdassan	wáccazdassan	kamyónaldassan	kamyónazdassan
Subessive Series				
Locative	wáccassul`	wáccazul`	kamyónall´ul`	kamyónazul`
Allative	wáccassul`e	wáccazul`e	kamyónall´ul`e	kamyónazul`e
Ablative	wáccassul`a	wáccazul`a	kamyónall´ul`a	kamyónazul`a
Translative	wáccassul`an	wáccazul`an	kamyónall´ul`an	kamyónazul`an
Apudessive Series				
Locative	wáccassuq	wáccazuq	kamyónall´uq	kamyónazuq
Allative	wáccassuqe	wáccazuqe	kamyónall´uqe	kamyónazuqe
Ablative	wáccassuqa	wáccazuqa	kamyónall´uqa	kamyónazuqa
Translative	wáccassuqan	wáccazuqan	kamyónall´uqan	kamyónazuqan
Inessive Series				
Locative	wáccassull´	wáccazull´	kamyónall´ull´	kamyónazull´
Allative	wáccassull´e	wáccazull´e	kamyónall´ull´e	kamyónazull´e
Ablative	wáccassull´a	wáccazull´a	kamyónall´ull´a	kamyónazull´a
Translative	wáccassull´an	wáccazull´an	kamyónall´ull´an	kamyónazull´an
	'brother'		'truck'	

not predictable from the root, though it does correlate with the ergative plural. The alternation between absolutive *wáccal* and ergative *wáccaz* in Table 13.7 is the only pattern that Charachidzé (1981: 45) describes as 'totally productive'. Other patterns apply to more restricted subclasses but remain nevertheless informative, as they allow a speaker to determine the plural grammatical cases from a single form.

The ergative forms also underlie the four series of secondary cases. In each series, the locative is based on the corresponding ergative and in turn underlies the other forms of its series. The last three series in Table 13.7 are particularly uniform. The locative is marked by *-ul`*, *-uq*, and *-ull`*, respectively, forming a base for an allative in *-e*, an ablative in *-a*, and a translative in *-an*. The superessive series departs only slightly from this pattern. In this series, the allative ending *-(d)e* alternates with the locative ending *-da*, and the ablative and translative endings *-ssa* and *-ssan* 'contain the segment *-ss-* (*-ssa*) which, under other conditions, indicates possession' (Charachidzé 1981: 45). Although the case inventory in Avar is larger than in Ingush,

and contains more complex forms, the uniformity of case markers determines similar kinds of mutually reinforcing patterns of interpredictability.

Given these patterns, a speaker can deduce the forms of an unfamiliar noun from the corresponding ergative, or from nearly any form based on the ergative. On a traditional analysis, for example, the ergative and genitive would be linked by mutual implication: if an ergative form of a noun is realized by *X* then the genitive is realized by *Xul*, and conversely. The same type of dependency holds between other pairs of forms. If one ablative form of a noun is realized by *X*, then the corresponding translative is realized by *Xn*, and vice versa. The network of dependencies that hold over a paradigm can be schematized as a set of implicational paradigm structure constraints of the form (ergative, *X*) $\approx$ (genitive, *Xul*). Or these patterns can be exhibited by sets of exemplary paradigms and extended to new forms by proportional analogies of the sort proposed initially by Paul (1920). Using the paradigm of *wacc* in Table 13.7 as an analogical base, one can deduce the nominative plural of *či* 'man', *číyal*, from the ergative plural *číyaz* by 'solving for *X*' in the proportion *wáccaz*: *wáccal* = *číyaz*: *X*.

As in Ingush, implications hold between paradigm cells rather than between forms in isolation. It is knowing that *wáccass* realizes the ergative singular that allows a speaker to identify the other singular grammatical cases, as well as the forms of the secondary series. The ergative can be regarded as the stem of noun paradigms in Avar, if 'stem' is interpreted to mean a recurrent element of form that is common to the members of a paradigm. The predictive value of a recurrent form is clearly exploited in schemas or proportions that use the ergative to deduce other case forms. However, the ergative is not uniquely informative, nor does its predictive value in analogical deductions force any particular morphotactic analysis.

If, however, one adopts a narrower construal of stems as 'building blocks', much of the flexibility of an implicational perspective is lost, because 'recurrent elements of form' must be construed as 'sub-word units'. A stem-based analysis in this 'derivational' sense would associate each Avar noun with a lexical stem set, including (among other forms) the ergative forms in the guise of 'oblique stems'. Ergative case would then be formally unmarked, and the grammatical cases based on the ergative forms would be taken to be built from the corresponding oblique stem. The distinction between inflectional stems and words becomes essentially distributional on this type of account; the ergative form corresponds to a stem because it underlies other forms of the paradigm, but the genitive is not a stem because it does not underlie any other forms. But what is the status of the locative forms? Since locatives underlie other forms of a series, they too would seem to be stems, formed from the oblique stem and what is sometimes termed a 'series marker'. What then about ablative forms, which appear to underlie the translative, even in the superessive series, where both forms contain the element *-ss-*? Do ablatives comprise yet another stem, which provides a base for a uniform translative

Table 13.8. The Super series in Lezgian (Haspelmath 1993a: §7.1.2)

Final Vowel	Ergative	Inessive	Superessive	Superdirective	Superelative	
Stressed	čarxú	čarxá	čarxál	čarxáldi	čarxálaj	'rock'
	čarčí	čarčé	čarčél	čarčéldi	čarčélaj	'paper'
Unstressed	nek'édi	nek'éda	nek'édal	nek'édaldi	nek'édilaj	'milk'
	šehérri	šehérra	šehérral	šehérraldi	šehérrilaj	'town'

marker *-n*? And if locatives and ablatives are not stems, but rather word forms that act as bases, on what principled grounds are ergatives treated as sub-word stems?

More generally, treating ergatives and other bases as sub-word units leads to a number of forced choices. A network of interpredictable forms may contain multiple implicational patterns that identify the realization of a given paradigm cell. But if this implicational structure is recast in derivational terms, as it must be in models that 'build' forms from an oblique stem, then it becomes necessary to select a particular derivation and choose a morphological analysis for the parts of a derived form. In many cases, there may be no principled basis for choosing between alternatives and no evidence that speakers are forced to make this kind of choice. The forms of the Lezgian Super series in Table 13.8 provide a further illustration.

Haspelmath (1993: 78) states that the inessive form 'is marked by lowering the final vowel of the oblique stem' (which realizes the ergative in Lezgian). This gives rise to the alternations between the final vowels in the ergative and inessive forms in Table 13.8. In the continuation of this passage, Haspelmath notes that corresponding superessive forms show the same vowel lowering, and thus can be derived from either the ergative or the inessive.

> The characteristic consonant of the Super localization is *-l*, which is added not directly to the oblique stem like the Ad, Sub and Post suffixes, but to a form whose final vowel has been lowered. Alternatively, one could say that the *-l* is attached to the Inessive case.

Haspelmath (p. 79) then turns to the relation of the superdirective to the superessive, and the formation of the superelative, which follows one of two stress-dependent patterns.

> The Superdirective is formed completely regularly from the Superessive by adding the Directive suffix *-di*, but a further complication arises in the Superelative case: Here the final vowel of the oblique stem suffix is lowered only when it is stressed, as shown in [Table 13.8].

From a traditional perspective, the Super cases exhibit multiple patterns of implication that permit a speaker to deduce new forms in more than one way. The superessive is identifiable from the inessive by a deduction that can be stated as (inessive, *X*) ≈ (superessive, *Xl*). Or exemplary ergative–superessive pairs can be used to deduce a new superessive X in a proportion such as *nek'édi*: *nek'édal* = *šehérri*: *X*.

The superdirective can likewise be deduced from the superessive, or from other exemplary pairs. The fact that only a superessive with a final stressed vowel predicts the vowel of the corresponding superelative can be incorporated into a deductive schema, or this generalization can be encapsulated in exemplary superessive–superelative pairs that exhibit the stress-conditioned variation described above. In short, a network of implicational relations can readily accommodate patterns in which a form *X* (e.g. *čarxú*) implies another form *Y* (e.g. *čarxá*), which in turn underlies a further form *Z* (e.g. *čarxál*), which itself provides the base for another form W (e.g. *čarxáldi*). Any one of the forms in the series *čarxá – čarxál – čarxáldi* implies the others, without dictating the status of either the stems that they share or of the endings that distinguish them.

However, an analysis that defines the Super series by adding exponents successively to an oblique stem is faced with a number of largely arbitrary choices. If the superessive is formed from the oblique stem, then this formation duplicates the vowel lowering in the inessive. But if the superessive is formed from the inessive, then either the analysis allows one case form to be derived from another – undermining the rationale for treating the ergative as a sub-word stem – or else the analysis is committed to treating the inessive also as a stem. The same issues arise in connection with the superdirective. If it is formed from the superessive, then either the analysis allows words to function as bases for other words, or the superessive is also a stem. If the superdirective is not formed from the superessive, then its formation must replicate the derivational steps that define the superessive base in *-l*. The ideal solution to these kinds of analytical choices is evidently a strategy that manages to avoid them altogether.

It is worth stressing that descriptive generalizations that make reference to notions like 'oblique stems' are wholly independent of the assumption that these stems are sub-word units. The rationale for treating stems as sub-word units is set out clearly in Kibrik's (1991: 257) discussion of the status of ergative/oblique stems in Daghestanian languages.

> Two different opinions can be found in the literature: (a) these markers are markers of the ergative case and all oblique cases are formed from the ergative; (b) these markers are markers of the oblique stem (of the singular or plural) and the ergative has no special marker and coincides with the oblique stem of the appropriate number. The first point of view is unsatisfactory: it does not take account of the semantics of the oblique cases (ergative meaning is not a component here), nor of the data from other Daghestanian languages, where the ergative frequently has a special morphological marker like other oblique cases . . .

The observation that 'ergative meaning is not a component' of the meaning of secondary cases merely shows that an ergative or genitive ***entry*** does not underlie the ***entries*** of secondary case forms in Archi or other Daghestanian languages. Only on a derivational account would one assume that stem syncretism in Archi involves a relation between entries rather than forms. From a traditional perspective, the

relation between these elements is purely formal: the form of the ergative implies the form of the corresponding secondary cases. This type of analysis also avoids the obscure 'semantics' that would be associated with an oblique stem. The observation that the ergative does not form a proper part of secondary case forms in some Daghestanian languages again counts against a derivational account on which secondary cases are built from the ergative. If the ergative form implies the form of the secondary cases, there is no reason that the ergative must be a proper part of the secondary cases.

In short, any insights stated in terms of oblique stems can be preserved by reinterpreting these stems as recurrent elements of form and excising the assumption that these elements must be sub-word units. This revision does not, of course, entail that recurrent units of form cannot be sub-word units. The fact that an ergative singular in *-í* implies an inessive in *-é* in Lezgian could be expressed by a schema (ergative, *Xí*) ≈ (inessive, *Xé*), in which the 'stem' *X* need not correspond to the nominative form, as in the case of pairs like *čarčí – čarčé*, in which *-čí* is the ergative marker. In the corresponding proportional analogy, which uses exemplary pairs such as *čarčí – čarčé* as an analogical base, the implicit stem is again the non-word *čarč*.

13.2.3.2 *Case cohorts*

The forms that make up a series in Avar and Lezgian exhibit a measure of morphosemantic coherence, and the same is true of local case series in other languages. However, sets of forms based on a common stem need not be morphosyntactically or morphosemantically coherent. Grade alternations in Sanskrit declensions provide one familiar example, in which differences in stem grade reflect the placement of Indo-European accent (Szemerényi 1990: 111ff). Finno-Ugric languages provide even more strikingly heterogeneous sets or 'cohorts' of forms based on a common stem. The paradigms in Table 13.9 below show the two patterns of gradation found

Table 13.9. Exemplary first declension paradigms in Saami (Bartens 1989: 511)

	PI ('Weakening')		PII ('Strengthening')	
	Singular	Plural	Singular	Plural
Nominative	**bihttá**	*bihtát*	*baste*	**basttet**
Genitive/Accusative	*bihtá*	*bihtáid*	**bastte**	**basttiid**
Illative	**bihttái**	*bihtáide*	*bastii*	**basttiide**
Locative	*bihtás*	*bihtáin*	**basttes**	**basttiin**
Comitative	*bihtáin*	*bihtáguin*	**basttiin**	**basttiiguin**
Essive	**bihttán**	**bihttán**	*basten*	*basten*
	'piece'		'spoon'	

978–0–19–920647–6 14-Spencer-ch13 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 216 of 218 May 16, 2008 18:58

in first declension nouns in (Northern) Saami. In the weakening pattern PI, the nominative and illative singulars and the essive forms are all strong, as indicated by consonant doubling in the forms *bihttá*, *bihttái* and *bihttán*. The remaining forms of the paradigm are weak in PI. The strengthening pattern PII distinguishes the same sets of forms, but inverts the strong/weak contrast. In PII, the nominative and illative singulars and the essive forms are weak (*baste*, *bastii*, and *basten*), and it is the other forms that are strong.

The forms of first declension nouns thus fall into the two heterogeneous sets. The nominative and illative singulars and the essive comprise one set, which is strong in PI and weak in PII. The other forms comprise a second set, which is weak in PI and strong in PII. As in a case series, each of the members of a cohort set predicts the form of the other members of the same set. But unlike locative series, neither of the sets in Saami can plausibly be characterized as a syntactic or semantic class. The division of these paradigms into purely formal cohort sets is reinforced by the syncretism between the comitative singular and the locative plural, two forms which again do not comprise any obvious natural class.

Estonian declensions present even more intricate patterns of interpredictability within sets of morphosyntactically heterogeneous forms. These patterns have been discussed in detail elsewhere (Viks 1992, Erelt et al. 2000, Blevins 2005, 2006), so a brief summary will suffice here. The paradigms in Table 13.10 show three of the patterns exhibited by first declension nouns. The paradigm of *maja* 'house' shows no grade alternations, but does contain an overlong short illative *`majja* whose first

Table 13.10. Exemplary first declension nouns in Estonian

	Singular	Plural	Singular	Plural	Singular	Plural
Nominative	**maja**	*majad*	**`tool**	*toolid*	**rida**	*`read*
Genitive	*maja*	**maja**de	*tooli*	**`tooli**de	*`rea*	**rida**de
Partitive	**maja**	**maja**sid	**`tooli**	**`tooli**sid	**rida**	**rida**sid
Stem Partitive		**maju**		**`toole**		**ridu**
Short Illative	**`majja**		**`tooli**		**`ritta**	
Illative	*majasse*	majadesse	*toolisse*	`toolidesse	*`reasse*	ridadesse
Inessive	*majas*	majades	*toolis*	`toolides	*`reas*	ridades
Elative	*majast*	majadest	*toolist*	`toolidest	*`reast*	ridadest
Allative	*majale*	majadele	*toolile*	`toolidele	*`reale*	ridadele
Adessive	*majal*	majadel	*toolil*	`toolidel	*`real*	ridadel
Ablative	*majalt*	majadelt	*toolilt*	`toolidelt	*`realt*	ridadelt
Translative	*majaks*	majadeks	*tooliks*	`toolideks	*`reaks*	ridadeks
Terminative	*majani*	majadeni	*toolini*	`toolideni	*`reani*	ridadeni
Essive	*majana*	majadena	*toolina*	`toolidena	*`reana*	ridadena
Abessive	*majata*	majadeta	*toolita*	`toolideta	*`reata*	ridadeta
Comitative	*majaga*	majadega	*tooliga*	`toolidega	*`reaga*	ridadega
	'house'		'chair'		'row'	

syllable is in the third quantity (Q3). (As in Viks 1992, Q3 syllables are marked with a preceding grave accent; see Lehiste (1997) for a recent summary of the phonetic properties of Q3 syllables.) The paradigms of *tool* 'chair' and *rida* 'row' both exhibit weakening gradation, in which the partitive singular and forms based on it are strong, whereas the genitive singular and the forms based on it are weak. In the paradigm of *tool*, grade is marked by the contrast between a Q3 partitive *`tooli* and a non-Q3 genitive *tooli*. In the paradigm of *rida*, the strong partitive *rida* is a disyllable and the weak genitive is an overlong monosyllable *`rea*, which is the result of a historical process of consonant elision. As these paradigms show, grade is an intrinsically morphological property, which correlates with phonetic quantity in different ways in distinct paradigms. Whereas *`majja* is grade-neutral, *`toole* is in the strong grade and *`rea* is in the weak grade.

As in Saami, the sets of interpredictable forms in Estonian are morphosyntactically heterogeneous. To trace just one path of dependencies, the partitive singular provides the base for the genitive plural, which in turn underlies each of the plural semantic cases, from the illative plural through the comitative plural. In a second cohort set, the genitive singular provides the base for the nominative plural and again underlies each of the singular semantic cases, from the illative through the comitative.

13.3 Conclusion

The paradigms described above highlight some recurrent patterns that arise in an especially clear form in relatively complex case systems. The discussion of analytical choices brings out a number of theoretical implications of these types of systems, and emphasizes the challenges that arise on analyses that attempt to 'build' the forms of complex case paradigms from sub-word stems. It is immaterial whether case forms are assembled in an 'item and arrangement' fashion (Hockett 1954) by suffixing morphemes to a stem, or whether they are derived in an 'item and process' manner by applying feature-adding rules to the stem. On either analysis, the multidimensional structure of a paradigm collapses to a set of independent derivations which are forced to assign often arbitrary morphotactic classifications to elements which are of less predictive value than the word forms from which they are extracted. By breaking a case paradigm down into a collection of independently-defined forms, a post-Bloomfieldian analysis obscures the network of relations that hold between forms.

Neo-Jakobsonian analyses that seek 'general meanings' for cohorts of case forms pursue a similar reductionist programme in a morphosyntactic domain. The appeal

to general meanings is one response to Kibrik's (1991: 257) objection that 'ergative meaning is not a component' of the meaning of secondary cases in Daghestanian languages. It is usually possible to describe Priscianic syncretism by classifying parasitic bases as stems that realize an 'abstract case meaning', though at the cost of introducing diacritic features with a purely distributional 'meaning'. This is particularly evident in Estonian, where the partitive meaning of `*tooli* 'chair' is not a component of the genitive plural meaning of `*toolide*, which in turn is not a component of plural semantic case meanings. An abstract stem corresponding to `*tooli* would thus be assigned features that it shares with the partitive singular and the genitive plural – as well as with the partitive plural `*toolisid* – but which contrast with the features of genitive singular *tooli*. A similar problem arises with the nominatives, given that the singular form `*tool* is based on the strong stem whereas the plural form *toolid* is based on the weak stem. One can set up a parallel inventory of 'stem features' to describe these patterns. However, the features will have no connection to genuine case 'meanings', general or specific. Instead, as in Stump (2001a), the abstract features will tag sets of forms with a common base, in the same way that the features in Jakobson (1936) tag forms with common endings.

Case paradigms, in short, represent a type of system 'où tout se tient' and it is only by treating paradigms as complex wholes that an analysis uncovers their organizing principles.

CHAPTER 14

CASE SYNCRETISM

MATTHEW BAERMAN

14.1 Introduction

By the term *case syncretism* we understand the combination of multiple distinct case values in a single form. Distinct case values are determined on a language-specific basis, so that case syncretism by this definition involves an observable asymmetry between paradigms within a language.[1] In the most obvious pattern, multiple case forms in one paradigm correspond to a single case form in another. For example, in Sanskrit, the singular paradigm of a-stem nouns has eight distinct case forms, which justifies positing eight distinct cases. In the non-singular paradigms, however, many of these distinctions are collapsed, particularly in the dual.

(1) Sanskrit a-stem noun (Whitney 1889: 114–15)

'god'	singular	dual	plural
NOM	devas	devāu	devās
VOC	deva	devāu	devās
ACC	devam	devāu	devān
INS	devena	devābhyām	Devāis
DAT	devāya	devābhyām	Devebhyas
ABL	devāt	devābhyām	Devebhyas
GEN	devasya	devayos	Devānām
LOC	deve	devayos	Deveṣu

[1] One might also identify an asymmetry along the diachronic dimension, where multiple cases in some older stage of a language correspond to a single case in a more recent stage (e.g. the various

Alternatively, distinct cases may be identifiable solely on distributional grounds (Zaliznjak 1973, Comrie 1991). For example, in Classical Armenian, the accusative never has a distinct form, being syncretic either with the nominative (singular nouns) or the locative (plural nouns, and pronouns).

(2) Classical Armenian (Meillet 1936: 81–2, 91)

	'father'	'fathers'	'I'
NOM	hayr	hark‘	es
ACC	hayr	hars	is
LOC	hawr	hars	is
GEN	hawr	harc‘	im
DAT	hawr	harc‘	inj
ABL	hawrē	harc‘	inēn
INS	harb	harbk‘	inew

Some apparent examples of case syncretism are doubtlessly superficial, and are best treated as accidental homophony. Thus in Latvian (Indo-European, Baltic), some declension classes have a nominative singular ending *-s*, and some have a genitive singular *-s*. The two sets overlap, so that there are declension classes which have both nominative and genitive singular in *-s*, resulting in syncretism of these two cases. However, since the *s*-endings occur independently of each other, there is no compelling reason to posit a systematic relationship between the two identical forms.

(3) Latvian (Forssman 2001: 113, 114, 118)

	a. i-stem 'brother'	b. a-stem 'hand'	c. u-stem 'market'
NOM SG	brālis	roka	tirgus
GEN SG	brāļa	rokas	tirgus

However, there is also considerable evidence that case syncretism may be systematic. For example, in Sanskrit nouns, the ablative and genitive singular are distinct only for a-stems (see *devas* above (1)). For all other noun classes they are syncretic, and the actual form may vary, as illustrated in (4).

oblique cases in (1) correspond to a single oblique case in many modern Indic languages). In this chapter, we restrict our focus to synchronically identifiable phenomena.

978–0–19–920647–6 15-Spencer-ch14 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 221 of 230 May 16, 2008 18:58

(4) Sanskrit (Whitney 1889: 119, 139)

	i-stem 'fire'	ṛ-stem 'giver'
NOM	agnis	dātā
VOC	agne	dātar
ACC	agnim	dātāram
INS	agninā	dātrā
DAT	agnaye	dātre
ABL	agnes	dātur
GEN	agnes	dātur
LOC	agnāu	dātari

Here it is clear that the syncretism is not accidental, as it is repeated across multiple exponents: the i-stems have the ending *-es*, while the ṛ-stems have *-u-* before the stem-final consonant.

14.2 Possible interpretations

Identifying a syncretic pattern as systematic still leaves open the question of which component of grammar encodes it, morphosyntax or morphology. On the first interpretation, syncretic case forms reflect some unity at the level of feature structure, an approach which owes a particular debt to Jakobson (1936). Thus, we might suppose that individual case values were in fact bundles of more abstract values. For example, we might say that 'ablative' is really a cover term for the feature bundle [+x, −y], while 'genitive' is [+x, +y]. Syncretism occurs in paradigms where only the shared value [+x] is active.

(5) Possible featural decomposition

	a-stem	i-stem, ṛ-stem
ABL [+x, +y]	[+x, +y]	[+x]
GEN [+x, −y]	[+x, −y]	[+x]

Viewed in this way, case syncretism is a window into the otherwise covert internal structure of case features. The second approach is to treat syncretism as a purely morphological phenomenon: the form of the ablative and genitive are stipulated as being identical, without reference to their function. On this interpretation, syncretism demonstrates the independence of morphosyntactic and morphological structure, and hence the autonomy of morphology.

For any individual language, it is hard, if not impossible, to fully justify one or the other approach on a principled basis. However, from a broader cross-linguistic perspective, each approach makes certain predictions. If case syncretism is motivated by shared meaning or function, then we should be able to predict which cases should be syncretic with each other, and find these repeated across languages. If, on the other hand, case syncretism is purely morphological, then we would not be able to make any cross-linguistic generalizations in terms of case function. As one might expect, the truth appears to lie somewhere in between.

14.3 Typology

The tendencies in case syncretism are clearest if we divide case syncretism into four types, numbered in decreasing order of their cross-linguistic frequencies. The figures are based on Baerman, Brown, and Corbett (2005), which is a survey of a controlled sample of 199 languages. Of these, seventy-four display inflectional marking of case on nominals. Forty of these languages have some sort of case syncretism: thus case syncretism is found in roughly half of the languages with inflectional case marking.

- Type 1: syncretism of the core cases, i.e. the cases expressing the grammatical roles of subject (transitive and intransitive) and object. This is found in thirty out of the forty languages with case syncretism (of the remaining ten, two do not distinguish the core cases inflectionally at all).
- Type 2: syncretism of a core case with some non-core case. Typically it involves what some might call the 'marked' core case (accusative or ergative, as opposed to nominative or absolutive). This is found in twenty-seven of the forty languages, often alongside type 1.
- Type 3: syncretism between non-core cases. This is found in five languages out of the forty, always alongside type 1 or 2.
- Type 4: case syncretism compounded with some other feature, such as number, found in five languages in the sample. These languages all display type 1 or 2 as well.

While the figures must be treated with some caution (since the identification of case syncretism is open to some individual interpretation), they do suggest that syncretism involving core cases predominates. Indeed, the presence of case syncretism in a language implies the presence of syncretism involving the core cases somewhere in the system.

14.3.1 Type 1

An example of nominative/accusative syncretism is given above in (2). Example (6) illustrates absolutive/ergative syncretism in Basque, which affects the proximate plural of all nouns (as well as the plural of some pronouns, e.g. *zuek* 'you.PL.ABS/ERG').

(6) Definite plural in Basque (Hualde 2003: 173)

	general plural 'the places'	proximate plural 'the places here'
ABS	lekuak	lekuok
ERG	lekuek	lekuok
DAT	lekuei	lekuoi

The cross-linguistic frequency of type 1 is perhaps not surprising when we consider it within the broader context of semantic and syntactic conditions that affect the marking of core arguments. For example, the notion of differential object marking describes the tendency for languages to restrict overt object marking (accusative case) to certain arguments, while conversely, the notion of differential subject marking describes how languages restrict overt transitive subject marking (ergative case). Both tendencies can be seen in Dhangu (Pama-Nyungan, Yolngu) where the accusative is marked on personal pronouns and the ergative is marked on inanimate nouns; animate nouns (in particular, human) fall in-between, marking both cases. We can group the two tendencies together under the rubric 'differential argument marking'.

(7) Core case syncretism in Dhangu (Schebeck 1976: 517, 526–7)

	personal pronoun 'we.DU. EXCL'	animate noun 'woman'	inanimate noun 'story'
ERG-INS	ŋalin^{y}u	taykkayu	ṯ̱āwuyu
ABS	ŋalin^{y}u	taykka	ṯ̱āwu
ACC	ŋalin^{y}un^{y}a	taykkan^{y}a	ṯ̱āwu

The idea behind differential argument marking is that there are prototypical patients and prototypical agents, the criteria being semantic or pragmatic, prototypical objects being inanimate, indefinite, and non-topical, prototypical agents being animate, definite, and topical. However, it is often the case that purely morphological or lexical categories play a role as well. For example, in Dhangu, although animacy plays the key role, there is a sharp distinction between pronouns and nouns, regardless of animacy. In Luiseño (Uto-Aztecan, Takic), inanimate nouns

lack the accusative ending that animate nouns have, in keeping with differential object marking. But an attributive adjective does mark the accusative, regardless of the animacy of the noun it is modifying.

(8) Luiseño (Malécot 1963: 208)

	animate		inanimate	
	man + bad = 'bad man'		grass + bad = 'bad grass'	
NOM	jaʔáš	alaxwɪš	šamʊt	alaxwɪš
ACC	jaʔátšɪ	alaxwɪtšɪ	šamʊt	alaxwɪtšɪ

More extreme examples of the morphologization of such patterns are familiar from Indo-European languages. Thus, in Latin, nominative and accusative are not distinguished by neuter nouns. While this would seem to follow from the notion of differential object marking (neuters are inanimates, and thus prototypical patients), the relationship between animacy and gender is not direct: inanimates are also well represented among the masculine and feminine genders, which do distinguish the two cases.

14.3.2 Type 2

In a nominative–accusative system, type 2 typically involves syncretism of the accusative with some non-core case, while in an ergative–absolutive system, it is the ergative which is involved. In Martuthunira (Pama-Nyungan), the accusative and genitive are syncretic for all nominals except vowel-stem nouns.

(9) Martuthunira (Dench 1995: 65, 107)

	V-stem 'money, rock'	N-stem 'tree'	L-/rr-stem 'daughter'	pronoun 'who'
NOM	kanyara	kalyaran	kurntal	ngana
ACC	kanyaraa	kalyaranku	kurntalyu	nganangu
GEN	kanyarawu	kalyaranku	kurntalyu	nganangu
LOC	kanyarala	kalyaranta	kurntala	nganala

In Araona (Tacanan, spoken in Bolivia), ergative and genitive are syncretic for all nominals except singular pronouns (nouns behave like the plural pronouns).

(10) Pronouns in Araona (Pitman 1980: 82)

	1SG	1DU	1PL
ABS	ema	tseda	cuama
ERG	yama	tseada	cuama(-ja)
GEN	quima	tseada	cuama(-ja)

Often the syncretic form appears to be an extension of the range of the non-core case form. For example, in Chukchi (Chukotko-Kamchatkan), the ergative (-instrumental) and locative cases are syncretic in the plural, and the form itself appears to result from an extension of the locative ending *-k* to the ergative. (The suffix *-r(ə)-* marks the oblique plural; the schwa is epenthetic–Skorik 1961: 204).

(11) Koryak third declension noun (Skorik 1961: 203–4)

'child'	singular	plural
ABS	nenenə	nenene-t
ERG-INS	nenene-te	nenene-rə-k
LOC	nenene-k	nenene-rə-k
ABL	nanana-jpə	nanana-r-ɣ əpə
DAT	nanana-ɣtə	nanana-rə-kə
ORI	nenene-ɣjit	nenene-rə-ɣjit
COM	ɣa-nanana-ma	ɣa-nanana-rə-ma

This may be combined with other patterns of syncretism yielding a situation where the marked core case has no dedicated form, and is only identifiable on distributional grounds. For example, in (12) and (13), type 2 syncretism is in complementary distribution with type 1 syncretism, in accusative and ergative systems, respectively.

(12) Estonian

'book'	singular	plural
NOM	raamat	raamatud
ACC	raamatu	raamatud
GEN	raamatu	raamatute
PART	raamatut	raamatuid

(13) Lak (Žirkov 1955: 36, 64–6)

	noun 'house'	pronoun'I'
ABS	k'atta	na
ERG	k'atlu-l	na
GEN	k'atlu-l	ttu-l
DAT	k'atlu-n	ttu-n

Such examples might be interpreted as a kind of differential argument marking, with the proviso that there is no dedicated form for the marked core case. Not all examples are of this type, though: for example, in Bonan (Mongolic), the accusative is always distinct from the nominative; however, the accusative itself is always syncretic with some non-core cases, depending on word class (with the genitive for nouns, and with the dative-locative for pronouns).

(14) Bonan (Todaeva 1997: 35)

	noun 'foliage'	pronoun 'he'
NOM	labčoŋ	ndžaŋ
GEN	labčoŋ-ne	ndžaŋ-ne
ACC	labčoŋ-ne	ndžaŋ-de
DAT-LOC	labčoŋ-de	ndžaŋ-de
ABL	labčoŋ-se	ndžaŋ-se
INS-COM	labčoŋ-*γ*ale	ndžaŋ-*γ*ale

Of course, for Bonan, an alternative analysis is possible: one might say that they reflect a quirky system of case assignment rules, whereby nominal direct objects were assigned genitive case and pronominal direct objects were assigned dative-locative. As always, there is a trade-off: the simpler the morphological rules, the more complicated the syntactic rules. In some cases, however, it is clear that the type 2 pattern is morphological and not syntactic. For example, in Russian, various classes of nominals display an alternation of type 1 and type 2 syncretism according to animacy: the accusative of inanimates is identical to the nominative, while the accusative of animates is identical to the genitive. The pattern is mirrored on agreeing adjectives.

(15) Inanimate ∼ animate alternation in Russian

	inanimate 'old chair'	animate 'old person'
NOM	staryj stol	staryj čelovek
ACC	staryj stol	starogo čeloveka
GEN	starogo stola	starogo čeloveka

However, this alternation is partly determined by inflectional class. This is evident when we contrast the behaviour of a-stem nouns with that of agreeing adjectives. The a-stems are an inflectional class which has the non-syncretic accusative case ending *-u* in the singular. The use of this ending is not sensitive to animacy. But if the noun is an animate masculine, the adjective will still display the syncretic pattern illustrated in (15):

(16) Animate a-stem with adjective

	'old man'	
NOM	staryj	mužčina
ACC	starogo	mužčinu
GEN	starogo	mužčiny

Thus, the pattern is morphological, not syntactic: although the adjective takes the form of the genitive, it is clearly functioning as an accusative, as evidenced by the

978–0–19–920647–6 15-Spencer-ch14 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 227 of 230 May 16, 2008 18:58

case marking on the noun. This is similar to what is found in the Luiseño paradigm in (8), where only the noun, but not the adjective, displays animacy-conditioned syncretism.

14.3.3 Type 3

Given the greater heterogeneity of the non-core case distinctions across different languages, one should not expect to be able to make precise comparisons. Nevertheless, some general patterns can be found. One important observation concerns the sort of exuberant patterns found in Indo-European, such as those found in Sanskrit (see (1) above). Though one might be tempted to see this as a reflection of an underlying network of morphosyntactic relationships, the fact is that these patterns are practically limited to Indo-European – unlike type 1 and type 2 syncretism, which are widely found in other language families. Isolated examples do occur, for example in Erzja Mordvin the dative and illative are syncretic in the singular of definite nouns. (That these two functions should be combined is perhaps unsurprising; however, that it should be limited to the definite singular forms is curious.)

One pattern which Indo-European *does* share with non-Indo-European languages is the collapse of all non-core cases, or of all non-nominative (non-absolutive) into a single 'oblique' form. Compare the paradigm of feminine nouns in Sogdian (Middle Iranian) in (17), with the Georgian attributive demonstrative in (18a).

(17) Sogdian (Sims-Williams 1982: 67)

	masculine 'person'	feminine 'tree'
NOM	ramí	waná
ACC	ramú	waná
GEN	ramé	wanyá
LOC	ramyá	wanyá
ABL	ramá	wanyá

(18) Georgian (Hewitt 1996: 97–8)

	a. 'this tree'		b. 'this one'
NOM	es	xe	es
ERG	am	xem	aman
DAT	am	xes	amas
GEN	am	xis	amis
INS	am	xit	amit
ADV	am	xed	amad

Note that, in spite of the superficial similarity, the nature of the patterns is quite different. In Sogdian, the syncretism is a consequence of the syncretic ending *-ya*. In Georgian, the syncretism is a consequence of the *lack* of an ending. The paradigm of attributive demonstratives in prenominal position is a reduced version of the full paradigm. In other contexts, e.g. as a free-standing pronoun (18b), the demonstrative shows full case inflection, alongside a stem alternation between the nominative and the other cases. In prenominal position, it retains the stem alternation but loses the case endings.

14.3.4 Type 4

Type 4 is a grab-bag of different patterns. A relatively simple example comes from Pashto (Iranian), which distinguishes two case forms (direct versus oblique) and two numbers (singular and plural). For some noun classes, the oblique singular is identical to the direct plural.

(19) Pashto (Grjunberg and Èdel'man 1987: 49–50)

	'person'		'farmer'	
	singular	plural	singular	plural
direct	saṛay	saṛi	bazgar	bazgər
oblique	saṛi	saṛio	bazgər	bazgaro

There have been some attempts to find a systematic principle in these patterns based on markedness relations between the cumulated morphosyntactic values (e.g. Serzisko 1982). For example, we might say that *saṛi* always realizes markedness disharmony between the case and number values: marked case (oblique) in the unmarked number (singular) or unmarked case (direct) in the marked number (plural). However, there is no consistent evidence to support this interpretation. Indeed, there are examples such as the comitative singular/locative plural syncretism of Saami (Hansson 2007; see below) which hardly lend themselves to a coherent analysis under anyone's conception of markedness, which suggests that this is the wrong line of inquiry.

14.4 Diachrony

Rather than seek some kind of unified explanation of syncretism, it may well be more fruitful to see it as the product of various kinds of diachronic processes. The data suggest that we need to consider at least five processes. The most widespread

type of case syncretism, that of the core cases, may in many instances represent the outcome of **desyntacticization**, that is, the morphologized relic of what was once an active syntactic rule. Thus, differential argument marking, in its purest form, represents a syntactic rule conditioned by semantic factors. The more morphological factors (such as inflection class) come into play, the less tenable a purely syntactic analysis becomes, to the point where one has a system where case choice is consistent across semantic classes, and it is morphology which determines patterns of formal identity. Of course, where to draw the line between a syntactic or morphological analysis of a given example may be a matter of some controversy.

Accidental syncretism occurs as a by-product of phonological or morphological changes. For example, the collapse of case forms due to phonological reduction or loss is a familiar theme in the history of Indo-European languages, as in Middle High German, where the reduction of unstressed vowels in final syllables to schwa (graphically *e*) in Middle High German led to widespread mergers throughout the inflectional system: compare Old High German *boto* NOM.SG ∼ *botun* ACC.SG ∼ *botin* GEN/DAT.SG 'messenger' to Middle High German *bote* NOM.SG ∼ *boten* 'ACC/GEN/DAT.SG' (Paul, Wiehl, and Grosse 1989: 201). The accidental effects of a morphological change are apparent in Russian, where the feminine instrumental singular ending *-oju* has become *-oj* over the last few centuries (which was not a regular phonological development). In adjectives and pronouns, where the ending *-oj* is also found for other cases, this led to syncretism, while in nouns it did not.

(20) Change in the instrumental singular in Russian

	adjective 'old'	noun 'book'
NOM	staraja	kniga
ACC	staruju	knigu
GEN	staroj	knigi
DAT	staroj	knige
LOC	staroj	knige
INS	staroju → staroj	knigoju → knigoj

That is, a morphological change affecting the shape of the ending led to syncretism in adjectives and pronouns, but the syncretism itself was not the motivating factor.

However, accidental patterns of syncretism may well be interpreted by language users as systematic, as evidenced by their **analogical extension**. For example, polysyllabic nominals in Saami regularly display syncretism of the comitative singular and locative plural, a pattern which clearly has no functional motivation, and seems to have been the accidental result of purely phonological developments. In most dialects, these forms are distinct for pronouns. However, in some northern dialects, this purely formal pattern of syncretism has been extended by analogy to monosyllabic nominals, through extension of the comitative singular form to the locative plural (Hansson 2007).

978–0–19–920647–6 15-Spencer-ch14 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 230 of 230 May 16, 2008 18:58

Alternatively, syncretism may arise from changes in the treatment of **case functions**. For example, in the early stages of Middle Indic (which descended from a system like that of Sanskrit), the genitive form had replaced the dative form in all paradigms except the singular of a-stem nouns (Pischel 1981: 289). This syncretism was not a result of phonological change; rather, there was a reassignment of case functions. The tasks originally allotted to the dative were reassigned to the genitive in most paradigms. Such reanalyses may ultimately lead to a shrinking or disappearance of case distinctions, as seen, for example, in many modern Indo-European languages (see fn. 1).

While the preceding three types involve the merger of originally distinct forms, syncretism can also conceivably result from the **partial bifurcation** of what was originally a single form. This is the scenario that Dench (1998) reconstructs for Martuthunira (see (9) above). The ending *-*ku* originally served for a single accusative-genitive case. In its role as a genitive, it could be followed by other case endings (double case or case stacking), which created a distinct set of phonological and morphotactic conditions which allowed *-*ku* to develop separately in the two case environments, at least for vowel-stem nouns. As a result, we can now speak of two distinct cases in Martuthunira, accusative and genitive. Where they are syncretic, this continues the original, undifferentiated state of affairs.

14.5 Conclusion

In principle, case syncretism may represent either of two things: (i) an underlying semantic or functional unity, such that the syncretic case form can be said to realize what is, at some level, a single case, or (ii) a morphologically defined identity of form. Clear examples of (i) may be taken to demonstrate interrelationships between the semantic or functional components of case, while clear examples of (ii) may be taken to demonstrate the autonomy of morphology. In practice, it is frequently difficult to make conclusive argument for either interpretation. Given that diachronic shifts between the two undoubtedly occur (as sketched in §14.4), this is not surprising. What this means is that not only is it important to recognize the dual significance of case syncretism, it is also important to remain circumspect in one's interpretation of it.

CHAPTER 15

THE DISTRIBUTION OF CASE

EDITH A. MORAVCSIK

15.1 Introduction

A comprehensive definition of case markers is as follows:

(1) A case marker is a formal device associated with a noun phrase that signals the grammatical role of that noun phrase.

Let us clarify the terms used in this definition. 'Formal device' includes segmental morphemes (affixes, clitics, stem modification, suppletion) and suprasegmentals (stress, pitch). Of these, the discussion in this paper will be restricted to segmental markers: case affixes and adpositions.[1] The term 'noun phrase' refers both to constituents overtly labelled as such in constituency grammars and to unlabelled units consisting of noun heads and their dependents in dependency grammars. 'Grammatical role' refers to the morphosyntactic reflections of semantic roles, such as subject and direct object.

The aspect of case marking that forms the topic of this article is its distribution. 'Distribution' refers to the conditions under which something must or may occur.

[1] Whether adpositions are properly termed case markers or whether the term should be restricted to affixes is a debated issue in the literature (cf. Chapter 12 and 33 in this volume). This paper adopts the broader concept.

Describing the distribution of an element within a linguistic structure involves two components: selection (what element occurs with what other elements) and linear order (temporal precedence relations among co-occurring elements).

As an initial working assumption, we will hypothesize that the distribution of case markers is simple.[2] If so, certain predictions follow. Let us first consider selection.

Selection

(A) **All case markers** are associated with noun phrases that have grammatical roles.
(B) **All noun phrases** that have grammatical roles are associated with case markers.
(C) In a given structure, each grammatical role is expressed by **not more than one case marker**, and each case marker expresses **not more than one grammatical role**.
(D) In a given language, **different grammatical roles** are expressed by **different case markers**, and **same grammatical roles** are expressed with the **same case markers**.

These four statements jointly posit bi-unique, one-to-one relations between grammatical roles and case marking. The statements in (A) and (B) exclude one-to-zero relations; those in (C) and (D) ban one-to-many relations.

The diagrams in (2) show what (A) and (B) require and what they exclude. (NP_{GR} stands for a noun phrase with a grammatical role; CM stands for case marker; lines indicate association.)

(2)

Required by (A) and (B)	Excluded by (A)	Excluded by (B)
NP_{GR} – CM	NP_{GR} – no CM	not a NP_{GR} – CM

Both (A) and (B) are empirical hypotheses: they may turn out to be true or false. The statement in (A) may seem logically implied by the definition of case markers in (1) but this is not so. By the definition, an element that marks **only** constituents other than noun phrases with grammatical roles would not qualify as a case marker, but the definition does allow for the possibility that a bona fide case marker's distribution is extended to constituents of other types; and this is what (A) excludes. The statement in (B) is also open to disconfirmation: the definition in (1) allows for the existence of noun phrases with grammatical roles that are not case-marked, which is what (B) bans.

The statements in (C) and (D) guard against one-to-many relations between individual grammatical roles and individual case markers syntagmatically and paradigmatically. This is shown in (3) and (4).

[2] For a similar approach to describing another cross-linguistically variant grammatical pattern: agreement, see Corbett 2006, esp. 8–27.

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 233 of 245 May 16, 2008 18:58

(3)

Required by (C)	Excluded by (C)
$GR_1 - CM_1$	$GR_1 - CM_1$ and CM_2 or GR_1 and $GR_2 - CM_1$

(4)

Required by (D)	Excluded by (D)
$GR_1 - CM_1$ $GR_2 - CM_2$	$GR_1 - CM_1$ \ CM_2 or GR_1 / $- CM_1$ GR_2

Statement (C) has to do with syntagmatic relations: in a given structure, there has to be exactly one grammatical role per one case marker and vice versa. Like those in (A) and (B), this pattern is not dictated by the definition of case markers in (1): noun phrases carrying multiple markings for a single grammatical role in a structure, excluded by (C), would be consistent with the definition. The statement in (D) has to do with paradigmatic relations: a given grammatical role having only a single possible case marker in a language, and distinct grammatical roles having distinct case markers. It thus excludes synonymous and homonymous case markers. Once again, departures from this stipulation are consistent with the definition in (1).

Let us now turn to the linear order of case markers relative to their hosts: what might be the simplest patterns?

Linear Order

(E) Every case marker has a **single allowable position** relative to the noun phrase that it marks.
(F) Every case marker is either **adjacent** to the noun phrase it marks or is contained within that noun phrase.

Although these are plausible patterns, they, too, may turn out to be false. Statement (E) excludes instances where a given case marker has variable positions. Its plausibility stems from the fact that segmental case markers are either affixes or clitics and such phonologically dependent formatives generally show invariant order. Statement (F) predicts iconic ordering widely documented in all aspects of morphology and syntax across languages: that parts of the sentence that go together stand together.

These considerations show that the six hypotheses above are vulnerable to evidence. In sections 15.2 and 15.3, we will take them up in turn and present some facts to assess their empirical adequacy. As we will see, none of them hold without qualifications.

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 234 of 245 May 16, 2008 18:58

15.2 Selection

We begin with the first of the two hypotheses that exclude one-to-zero relations between case markers and noun phrases with grammatical roles.

(A) **ALL** CASE MARKERS ARE ASSOCIATED WITH NOUN PHRASES THAT HAVE GRAMMATICAL ROLES.

This statement is generally true: case markers occur on such noun phrases as subjects, objects, and other complements and adjuncts, determiners, and modifiers of verbs, adjectives, adpositions, and nouns. Examples are given in (5).

(5) (a) case-marking of the object of a verb
*John has met **him**.*
(b) case-marking of the complement of an adjective
*John is jealous **of him**.*
(c) case-marking of the complement of a preposition
*I heard this from **him**.*
(d) case-marking of the determiner of a noun
***John's** father is retired.*
(e) case-marking of the modifier of a noun
*The letter **from him** was delayed.*

There are nonetheless occurrences of case markers that deviate from this pattern. Examples are in (6).

(6) (a) ***To** your health!*
(b) *Now, **about** your grade.*
(c) German
Den *rot-**en**.*
the:ACC red-ACC
'The red one.' (in answer to questions like 'Which one do you want?')
(d) Latin
Me *miser-**um**!*
I-ACC miserable-ACC
'Oh, unhappy me!' (Blake 1994: 9)
(e) Serbo-Croatian
*Gospodin-**e**!*
gentleman-VOC
'Sir!'
(f) Latin
Hōc *respons-**ō*** *dat-**ō*** *discessit.*
this:ABL response-ABL given-ABL he:departed
'This answer given, he departed.'
(g) ***With** the problem safely out of the way, we can move on.*

In the first five examples, the sentences consist of single noun phrases which therefore cannot possibly bear grammatical roles relative to another constituent. In (6a)–(6c), case-marking is justified at least semantically since there are understood governors: 'let us drink' in (6a), 'let us talk' in (6b), and 'want' in (6c). But the exclamatory accusative in (6d), the vocative in (6e), the ablative absolute construction in (6f), and the absolute *with*-construction in (6g) have no similar semantic justification.[3]

In (6), case occurs on constituents that have no grammatical role; but the constituents are full noun phrases. A different departure from (A) is when constituents are case-marked that are not full-fledged noun phrases. An example is predicate complements (Blake 1994: 95–7; Spencer and Otoguro 2005: 122–3)

(7) (a) Latin
Cicerōn-em ***cōnsul-em*** *creāvērunt.*
Cicero-ACC **consul-ACC** made:3PL
'They made Cicero consul.' (Blake 1994: 95)

(b) Hungarian
Peter ***beteg-nek*** *látszik.*
Peter **sick-DAT** seems
'Peter seems sick.'

In the Latin example, the predicate complement is a noun in the accusative; in the Hungarian example, it is a nominalized adjective in the dative. These predicate complements are not prototypical noun phrases: for example, they cannot be referential.[4]

The second hypothesis excluding one-to-zero relations between case markers and noun phrases bearing grammatical roles is (B).

(B) **ALL** NOUN PHRASES THAT HAVE GRAMMATICAL ROLES ARE ASSOCIATED WITH CASE MARKERS.

Four departures from this simple pattern occur. First, whether a noun phrase is or is not marked for case may depend on the case involved. Certain cases are particularly prone to being unmarked. Lehmann (1985) proposes the following:

(8) Given the scale of Nominative < Accusative < Genitive < other cases, if a language has zero-marking for a case, it also has zero-marking for all cases to the left.

(On the issue of zero-marking versus no marking, see sections 12.1 and 12.6 in this volume.)

[3] On the vocative, see Spencer and Otoguro 2005: 129–33.

[4] The distribution of case markers may extend even beyond non-prototypical noun phrases. The same forms that mark adpositional case on noun phrases may also occur as prefixes as in the English word ***in****come* or in Serbo-Croatian ***po****duzimati* 'to undertake'. Additional Slavic and other European languages provide many other examples.

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 236 of 245 May 16, 2008 18:58

Second, the occurrence of case-marking may vary with the type of noun phrase: for example, pronouns are more commonly case-marked than noun-headed phrases (cf. Iggesen 2005a and Chapter 16 in this volume); definite noun phrases are more commonly accusative-marked than indefinite ones (cf. Chapter 22 in this volume), and infinitival and clausal noun phrases are least likely to bear case-marking. (On variable case marking depending on the type of noun phrase, see also sections 4.1 and 6.3 of Chapter 12 in this volume.)

Third, the presence of case marking may vary depending on whether it is needed for clarity. In Hungarian, the genitive case suffix is obligatory only if the possessor deviates from its usual order by following, rather than preceding, the possessum, or if it precedes but is non-adjacent to it. If it immediately precedes the possessum, its use is optional. Case markers may also be redundant and thus dispensable if the grammatical role of a noun is rendered likely by its lexical meaning, as in *I will see you* ***(on) Tuesday.*** On other instances of optional case marking, see Chapter 22 in this volume; on case-marker-less locative phrases, see Chapter 46.

The fourth exception to (B) has to do with instances where compliance would lead to a sequence of more than one case marker on a single nominal. In some instances, spelling out both case markers is grammatical, as in (9).

(9) (a) Quechua (Blake 1994: 104)
*Hwan-**pa-a** rikaa.*
John-GEN-ACC I:see
'I see John's (house).'

(b) Basque (Moravcsik 2003: 453)
*gizon-**aren-ari***
man-GEN-DAT
'to that of the man'

At first blush, the word *Hwan-pa-a* 'John's' in (9a) is paradoxically marked: it has both a genitive and an accusative suffix. The paradox disappears, however, if we realize that the word stands for two noun phrases with distinct roles. One is 'John', the possessor of 'house', which is understood; the other is 'John's house', which is the direct object of 'see'. The marking of the two grammatical roles results in a sequence of two case markers. The same holds for (9b). (10) provides a similar example from English.

(10) *I gave the picture to Anne's brother and the book* ***to*** *Susan's.*

In *to Susan's*, the two noun phrases involved are the understood 'the brother' marked for the dative case and 'Susan', the possessor of 'brother', marked for genitive.

This construction, known as hypostasis – consisting of a genitive determiner that lacks its head and is marked both for genitive and for the case assigned to the understood head (cf. Moravcsik 2003: 453–4) – makes perfect syntactic sense: it is

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 237 of 245 May 16, 2008 18:58

in compliance with (B), which requires that all noun phrases with grammatical roles should be marked for their roles. In spite of its reasonableness, multiple case marking of this kind is rare across languages, resulting in common violations of (B).

We now turn to the two hypotheses that ban one-to-many relations between case markers and grammatical roles. (C) does so syntagmatically.

(C) IN A GIVEN STRUCTURE, EACH GRAMMATICAL ROLE IS EXPRESSED BY **NOT MORE THAN ONE CASE MARKER**, AND EACH CASE MARKER EXPRESSES **NOT MORE THAN ONE GRAMMATICAL ROLE**.

This statement makes two claims. The second of the two is that a case marker does not express more than one grammatical role at the same time. This is generally true; a possible exception is the 'designative case' in Even, which single-handedly marks a noun as an object and its possessor as a beneficiary; see Chapter 44 in this volume.

The other claim is that a grammatical role is not expressed by more than one case marker in a given structure. This is indeed a common pattern. The single adposition or case affix may mark a grammatical role on the noun phrase as a whole as shown by the preposition in (11a), the postposition in (11b), and the noun-phrase-final affix in (11c).

(11) (a) English
with *the three young boys*

(b) Japanese[5]
Sono wakai otokonoko ***ga*** *furui pan* ***wo*** *tabeteiru.*
the young boy NOM old bread ACC eat
'The young boy is eating old bread.'

(c) Basque (Saltarelli et al. 1988: 77, 300)
*etxe zaharr-****etan***
house old-LOC
'in old house'

Another pattern is where the single case marker is associated with a proper subpart of the noun phrase, such as the noun or the demonstrative (Blake 1994: 99–106).

Alternatively, however, there may be more than one marker for a given case within the noun phrase. This departure from the one-noun-phrase–one-case-marker relation stipulated by (C) takes two forms: either more than one constituent of the noun phrase is marked by case; or the case marker is itself compounded, consisting of more than one marker of the same case.

First, let us consider extended exponence: instances where more than one constituent of a noun phrase is case-marked for the same grammatical role. Four of the

[5] For the multiple functions of Japanese postpositions, see Spencer and Otoguro 2005: 134–41 and Chapter 12 in this volume.

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 238 of 245 May 16, 2008 18:58

possible configurations of case concord are exemplified in (12) (for relevant cross-linguistic generalizations and discussion, see Corbett 2006: 133–5) and universals #119 and #120 in the Konstanz Universals Archive).

(12) (a) Hungarian: **DEM** ART ADJ **NOUN**
*ez-**t*** *a* *magas* *ember-**t***
this-ACC the tall man-**ACC**
'this tall man (ACC)'

(b) Duungidjawu: **NOUN ADJ** DEM (Wurm 1976: 108–9)
*d^{y}an-bam-**ma*** *būgubu-**na*** *man*
men-DUAL-**ACC** short-**ACC** this
'these two short men (ACC)'

(c) German: **DEM ADJ** NOUN
*diese-**n*** *große-**n*** *Mann*
this-**ACC** tall-**ACC** man
'this tall man (ACC)'

(d) Latin: **DEM NOUN ADJ**
hunc *vir-**um*** *magn-**um***
this.**ACC** man-**ACC** tall-**ACC**
'this large man (ACC)'

An unusual instance of case concord occurs in Romani, where adjectives agreeing with their nouns distinguish only two historically prior cases whereas nouns are marked for several cases (Koptjevskaja-Tamm 2000: 123). For similar mismatches in case concord in German and Estonian, see sections 12.6.2 and 12.6.8.2 in this volume.

Whether concord does or does not take place may depend not only on the particular constituents of the noun phrase, as shown in (12), but also on other factors. One is linear position: the same word may or may not agree depending on where it is located. As illustrated in (13), concord is preferred if that word is non-adjacent to the noun.

(13) Warlpiri (Hale 1973: 314; transcription simplified)
(a) *Tjantu* *wiri-**ngki*** *tji* *yalkunu.*
dog big-**ERG** me bit
'The big dog bit me.'
(b) *Tjantu-**ngku*** *tju* *yalkunu* *wiri-**ngki.***
dog-**ERG** me bit big-**ERG**
'The big dog bit me.'

In addition to the particular noun phrase constituent and its position within the phrase, two other factors play a role in whether concord does or does not take place: the meaning of the case and its form. First, there may be concord for the marking

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 239 of 245 May 16, 2008 18:58

of some cases but not for others. In Estonian, there is concord for the adjective and the noun in the elative case but not in the comitative (Kilby 1981: 115).

Second, affixes and adpositions are not equally prone to concordial case marking. Kilby (1981: 115, 118) observed that, in contrast with affixal case markers, which may show concord, adpositions occur only once per noun phrase. While this may be the prevailing pattern, there are instances of noun-phrase-internal agreement in adpositions. In Hungarian, the demonstrative optionally repeats the postposition of the entire noun phrase (for more detail, see Moravcsik 2003: 207–9).

(14) (a) *ez-**en*** *a* *nap-**on*** ***kívül***
this-SRESS the day-SRESS **besides**
'besides this day'
(b) *ez-**en*** ***kívül*** *a* *nap-on* ***kívül***
this-SRESS **besides** the day-SRESS **besides**
'besides this day'

Case concord may result in a special morphological type of construction known as Suffixaufnahme (cf. Plank 1995a; and Chapter 12, 44, and 52 in this volume). For an example, take a noun phrase that consists of a noun head and its possessor, such as 'the man's wife'. The entire noun phrase will have a particular grammatical role in the sentence and 'the man's' also bears a grammatical role: that of a possessor. If there is to be case concord in this noun phrase, 'the man's' will have to have two case markers: one for its own possessor relation and one for the case of the entire noun phrase. A pattern of double case-marking would result, as in (15).

(15) Kayardild (Evans 1995b: 398)
*thabuju-**karra-nguni*** *mijil-**nguni***
brother-GEN-INS net-INS
'with brother's net'

This sequence of two case markers on a single constituent is similar to that seen in the Quechua and Basque examples of hypostasis in (9). However, there is also a difference: in (9), the external case descends on the possessor constituent due to lack of an overt head, while here, there is an overt head.

Suffixaufnahme is syntactically entirely regular and thus one would expect it to be very common across languages. Since it is not common, the question is how languages manage to avoid it. Two obvious resolutions are: either the external case is suspended or the internal case is (cf. Blake 1994: 104–5). The two alternatives are illustrated in (16), respectively.

(16) (a) Latin
*alb-**am*** *columb-**am*** *puell-**ae***
white-ACC dove-ACC girl-GEN
'the girl's white dove'

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 240 of 245 May 16, 2008 18:58

(b) Classical Armenian (Blake 1994: 105)
bazmnt'eamb zarauk'm Hayoc
mass.INS troops.INS Armenian
'with the mass of Armenian troops'

In the Latin example, concord applies to the adjective but bypasses the genitive-marked constituent. In Armenian, it is the genitive that yields: 'troops' is expected to be in the genitive but instead it takes on the external case.

In addition to multi-constituental noun phrases with a single head, another construction type where a noun phrase may be multiply marked for its grammatical role is coordination. Consider (17).

(17) (a) *Students come **from** Germany and Italy.*
(b) *Students come **from** Germany and **from** Italy.*

In coordinate structures, it is the higher noun phrase whose case is governed by the verb or some other governor. Correspondingly, (17a) may be analysed as the case-marking preposition assigned to this higher noun phrase. In (17b), however, the higher noun phrase's case marking percolates to the two conjuncts.

As shown in (17), in English, such double case-marking of conjuncts is an option that co-exists with single marking. This is often so if the case marker is an adposition. If the case marker is an affix, separate marking of each conjunct is generally required as shown in (18) (but see Wälchli 2005, sections 2.2 and 3.2; Spencer and Otoguro 2005: 128; and section 12.6.8.1 in this volume on 'suspended affixation').

(18) Hungarian
(a) *két fiú-**t** és két lány-**t***
two boy-ACC and two girl-ACC
'two boys and two girls(ACC)'
két fiú és két lány-t***
two boy and two girl-ACC

In the examples discussed so far, the one-to-one relationship between noun phrase and case marker was departed from by **more than one constituent of a noun phrase marked for case**. The Hungarian examples in (14) already evidenced the other departure already mentioned above**: the case marker itself consisting of more than one marker**. The most common form of this is the one shown in (14), where an adposition and an affixal case jointly mark a grammatical role with the adposition itself governing an affixal case. The pattern is familiar from German, Russian, or Latin; (19) illustrates it from Ancient Greek.

(19) Ancient Greek
(a) ***eis** tē-**n** poli-**n***
into the-ACC city-ACC
'into the city'

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 241 of 245 May 16, 2008 18:58

(b) ***en*** *tē* *pol-**ei***
in the-**DAT** city-**DAT**
'in the city'

Alternatively, there may be more than one adposition – as in English ***from out of*** *two hundred candidates* – or more than one affix. The affixes may be the same: in Hungarian, the accusative affix may be doubled on some stems, e.g. *ö-t-et* 'he/she-**ACC-ACC**' 'him/her'. Or, a case marker may be composed of more than one case marker, as in Avar, where the suffix *q-e* 'to' consists of *q* 'at' and *e* 'toward' (cf. section 12.5 and Chapter of this volume). Compound case-markers are frequent in Australian languages (Blake 1994: 154).

The crucial property of the patterns discussed above that puts them at odds with the statement in (C) is over-marking, or redundancy. It should be noted that grammatical roles may be redundantly expressed not only by multiple case marking but also if in addition to case marking, other formal devices are also employed to identify them in a sentence, such as constituent order and verb agreement (see Chapter 19 in this volume). In many languages, grammatical roles may be signalled in all three of these ways in the same sentence, as in English *She visits him often*. While here case, verb agreement, and word order converge in identifying the subject, the three devices may signal different kinds of grammatical roles. A cross-linguistically recurrent mismatch is between case marking that follows the ergative pattern, and verb agreement based on the accusative alignment (cf. Corbett 2006: 56–8).

We will now turn to the statement excluding one-to-many relations between grammatical roles and case markers on the paradigmatic level.

(D) IN A GIVEN LANGUAGE, **DIFFERENT GRAMMATICAL ROLES** ARE EXPRESSED BY **DIFFERENT CASE MARKERS**, AND **SAME GRAMMATICAL ROLES** ARE EXPRESSED WITH THE **SAME CASE MARKERS**.

The first assertion contained in (D), that different grammatical roles are signalled by different case markers, is commonly departed from. Case syncretism violating this pattern is discussed in Chapter 14 of this volume; see also Baerman and Brown 2005.

The second claim of (D): same grammatical role – same case marker, might hold true if the assignment of case markers were dependent only on grammatical role. This, however, is not so for three reasons.

First, the form of the case marker frequently depends on the morphosyntactic class membership of the constituent that carries it (cf. section 12.2 in this volume). Nouns, adjectives, and pronouns may carry different markers for the same case, as in Russian:

(20) Russian
*iz bols'-**ovo** gorod-**a***
from large-GEN city-GEN
'from a/the large city'

Second, markers of the same case may vary not only with the diverging properties of the case-marked noun phrase but also with the divergent properties of the governing constituent. If the same grammatical role were associated with the same case markers, arguments of finite verbs and arguments of nominalized versions of those verbs would be marked the same way. This is not so: at least one of the arguments of a nominalized verb is generally expressed as a genitive.[6]

(21) Lithuanian (Koptjevskaja-Tamm 2003b: 733; glossing simplified)
(a) *Amerik-**a** buvo atrasta Kolumb-o*
America-NOM was discovered Columbus-GEN
'America was discovered by Columbus'
(b) *Kolumb-o Amerik-**os** atradimas*
Columbus-GEN America-GEN discovery
'Columbus' discovery of America'

But even finite verbs may call for different markers on their complements for what seems to be the same case depending on the verb's semantic and syntactic properties. High-level transitivity as defined by Hopper and Thompson favors the accusative marking of objects over oblique marking (Hopper and Thompson 1982). In some languages such as Russian and Finnish, the case marking of objects differs depending on whether the clause is affirmative or negative. Verbs referring to psychological states, verbs of perception, verbs with modal meanings, and verbs of possession generally show non-canonical marking of their subjects and objects (Onishi 2001; Aikhenvald, Dixon, and Onishi 2001; Chapter 22 in this volume; cf. also universal #123 in the Konstanz Universals Archive). (22) is an illustration contrasting the absolutive complement of the verb 'cut' and the instrumental complement of the verb 'want' in Samoan.

(22) Samoan
(a) *E sogi e le tama **O** le ufi i le to'i.*
PRES cut ERG the boy ABS the yam INS the knife
'The boy cuts the yam with the knife.'
(b) *E mana'o le tama **i** le teine.*
PRES want the boy INS the girl
'The boy wants the girl.'

[6] Nonetheless, it is not entirely clear whether the argument relations of a verb and of the nominalized form of the verb are the same. If they can be shown not to be the same, different case markings of arguments in the two constructions do not form a deviation from point (D). On this and related issues of nominalization, see Malchukov 2004.

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 243 of 245 May 16, 2008 18:58

Third, case may cumulate with other properties of the noun phrase that is marked for it, such as gender, number, or definiteness. This means that noun phrases with the same grammatical role may carry different case markers if they differ in these other properties. This pattern is well-known from Indo-European languages – such as Germanic, Slavic, and Romance – where case cumulates with gender and number (cf. Spencer and Otoguro 2005: 123–5 and sections 12.1 and 12.2 of this volume). The effect of definiteness on the case marking of direct objects is illustrated in (23): in Turkish, definite or specific objects are overtly marked while others have no marking.

(23) Turkish
  (a) *Kitablar-ı okumadım.*
  books-DEF.ACC I:have:not:read
  'I have not read the books.'
  (b) *Kitablar okumadım.*
  books:INDEF.ACC I:have:not:read
  'I have not read books.'

Let us now consider the linear relations of case markers and their noun phrases in their simplest forms and possible deviations from them.

15.3 Linear order

(E) EVERY CASE MARKER HAS A **SINGLE ALLOWABLE POSITION** RELATIVE TO THE NOUN PHRASE THAT IT MARKS.

Generally, case markers do indeed have unique positions relative to the noun phrase. There are nonetheless three types of departures from this pattern.

First, there may be freely variant order between noun phrase and adposition, such as in English ***notwithstanding*** *the weather*, and *the weather* ***notwithstanding***.

Second, there may be free variation in the ordering of affixes as in Zyryan (Comrie 1980: 81–2; see also section 12.5 of this volume).

(24) Zyryan
  (a) *kerka-**nim-lan***
  house-**our-to**
  'to our house'
  (b) *kerka-**lan-nim***
  house-**to-our**
  'to our house'

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 244 of 245 May 16, 2008 18:58

Third, the order of case marker and noun phrase may depend on the type of noun phrase. In many languages, the case marking of nouns is in terms of distinct morphemes while the case marking of pronouns is suppletive, which means that whatever is the order of stem and case marker for nouns, it does not hold for pronouns.

(F) EVERY CASE MARKER IS EITHER **ADJACENT** TO THE NOUN PHRASE IT MARKS OR IS CONTAINED WITHIN THAT NOUN PHRASE.

This statement covers the general pattern of adpositional order: adpositions are generally adjacent to their noun phrases or their components (on inpositions, see Dryer 2005b). English preposition-stranding – as in ***Who*** *are you talking* ***about***? – departs from this pattern. The construction, which is infrequent across languages, illustrates the less than fully clear-cut difference between case marker and verbal particle. For an Iraqw pattern which violates the adjacency requirement, see section 44.2.3 in this volume.

Statement (F) also predicts the prevalent order of affixal case markers relative to other affixes. As noted by Greenberg (1963b, universal #39), there is a cross-linguistic tendency for the number affix to come between the stem and the case affix. Bybee (1985: 33–4) relates this to a general relevance principle according to which affixes whose meaning is more closely relevant to that of the stem are ordered closer to the stem than those whose meaning is less relevant. Putting it differently, linear order reflects semantic scope relations: since case has scope over a noun including its number, it should be ordered more externally.

Nonetheless, (F) is not exceptionless. Of the two alternative orders in Zyryan shown in (24), only (a), where case is outside the possessive marker, conforms to the pattern. Several other such non-iconic affix orders in Uralic languages are discussed in Comrie 1980; for examples from Finnic and Finno-Volgaic languages, see Spencer 1992: 320–3 and from other languages, Blake 1994: 106.

The adjacency of noun phrase and case marker may be realized by different linear precedence relations: the case marker may immediately precede or immediately follow the noun phrase or its relevant parts. In this regard, however, there is an asymmetry between adpositions and case affixes. Adpositions are variable in their order: some languages have prepositions, others have postpositions, again others have both. The occurrence of the two orders correlates with the basic constituent order type of a language (Dryer 2005c). Affixal case markers in turn show a strong preference for following the constituent they occur on. In his language sample, Dryer has found 431 languages with core-case suffixes as opposed to only 35 with prefixes, such as Takelma (Dryer 2005b).

(25) Takelma
*'**gwel**-danà*
under-rock
'under the rock'

978–0–19–920647–6 16-Spencer-ch15 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 245 of 245 May 16, 2008 18:58

This fact is in line with the general preference for suffixing over prefixing of bound markers in general (Stump 2001b, Dryer 2005a).

15.4 On explanations

Several of the facts about the distribution of case marking surveyed in this paper have a functional basis. Thus, the overwhelmingly preferred adjacence of case marker and noun phrase, which is a consequence of the more general tendency for semantically coherent constituents to stand next to each other, clearly helps comprehension. The preference for the suffixing of case markers over their prefixing – a reflection of a cross-linguistic preference for suffixes over prefixes regardless of what they express – also aids the decoder. As suggested by Hawkins and Cutler (1988), 'speakers and listeners process stems before affixes' (306). If so, receiving stems first facilitates natural processing.

However, how function comes to shape form cannot be understood without reference to the historical processes that provide the mechanisms to mediate between the two. For the explanation of the suffixing preference, Hall proposed that word-initial position resists phonological fusion and thus the prefixation of lexical elements – the ancestors of case markers – to stems, while word-final position favours it (Hall 1984). Additional diachronic factors determining the evolution of case markers and their distribution are discussed in Part V of this volume.

Acknowledgement

I am very grateful for the detailed and insightful comments I received from the editors of this volume.

CHAPTER 16

ASYMMETRY IN CASE MARKING

NOMINAL VS. PRONOMINAL SYSTEMS

OLIVER A. IGGESEN

16.1 Introduction: the concept of case-asymmetry

Many languages with nominal case inflection apply the same distinctions of case forms evenly across their entire nominal lexicon; hence, all conceivable subclasses of nominals must unequivocally be analysed as exhibiting exactly the same inventory of case categories. This morphological state of affairs will henceforth be called **case-symmetry**. A straightforward example in point is Turkish, where the same suffixes code the same case functions with all nominals. In fusional languages like Russian, changing homophony patterns of case forms among the inflection classes (and within the word paradigms) require the application of the distributional method in order to identify their underlying category system. In terms of the range of lexemes over which these categories have scope, however, these languages can usually be classified as case-symmetrical as well.

There are, however, languages in which certain (or all) case distinctions apply rather selectively to only a subset of their nominals, in such a way that the inflectional paradigms of the minority subclass can be envisaged as containing more, less, or substantially different case categories than the bulk of the nominals. In other words, certain morphological cases arguably fail to have scope over the entire nominal lexicon in such languages. These languages will henceforth be referred to as **case-asymmetrical**.[1] The theoretical implications and argumentative rationale of this – purely morphology-based – analysis will be addressed in section 16.2, following a particularly well-known illustration of case-asymmetrical paradigm structure: the opposition of direct vs. objective case in English, which applies only to a small number of pronominal lexemes.

(1) English[2]

	full noun	1 **sg.**	3SG. M.	3 SG. F.	1 PL.	3 PL.
DIR	man	I	he	she	we	they
OBJ		me	him	her	us	them

To these five lexical items, the human interrogative-relative pronoun *who* (OBJ *whom*) must be added as well. No other noun or pronoun distinguishes an objective form in English. Similar constellations are also found in a number of genetically and areally related languages (e.g. Spanish, Swedish, and Bulgarian). Since there had been little awareness of case-asymmetry in unrelated languages until recently,[3] the European evidence was generally dismissed as typologically irrelevant (e.g. Nichols 1992, claiming what amounts to our concept of case-asymmetry to be a low-frequency phenomenon on the world-wide scale). In the 260-languages cross-linguistic survey by Iggesen (2005b), however, more than one quarter of the investigated languages were shown to be case-asymmetrical (whereas case-symmetrical languages made up more than a third). Furthermore, the study demonstrated that the locus of case-asymmetry is overwhelmingly the personal pronouns (more precisely, the SAP pronouns) rather than other conceivable subclasses of nominals (cf. the definition of legitimate loci of deviant case inventories in section 16.2 below). In view of the strong association of case-asymmetry with the pronouns, the present contribution is limited to evidence from this particular lexeme class.

[1] The concept and terminology of case-(a)symmetry have been introduced in Iggesen (2005b).

[2] In all tables in this chapter, single cells represent unified morphological categories, covering one or more of the case functions that are indicated in the leftmost column. The mismatch between the cells visualizes the extent and quality of the asymmetry constellation.

[3] With the notable exception of the paradigm structure type classified here as qualitative case-asymmetry (see below).

16.2 Theoretical issues and definitions

While there is no doubt that the languages deemed case-asymmetrical here are characterized by having a deviant set of distinct case expressions in a minority subset of nominals (for the current purposes, in the pronouns), it remains to be argued for whether this deviation can be equated with a different set of case categories (i.e. underlying paradigmatic oppositions). A rigid application of the distributional method would entail that all nominals of any given language make, overtly or covertly, exactly the same category distinctions.[4] This would mean that case-asymmetry does not exist as such, and the constellations referred to here as case-asymmetrical would be reduced to a special, lexeme-class-sensitive subtype of case-syncretism (cf. Baerman in this volume). However, many linguistic analyses, including the various notions of case, are descriptive models serving particular purposes; with respect to the (still imperfectly understood) cognitive reality of language, they can be – more or less adequate – approximations at best. Thus, a radically distributional approach has the advantage of being descriptively elegant and economical for the analysis of syntax, as it keeps the number of rules at a minimum. It can be counterintuitive and awkward when it comes to evaluating morphological regularities, though. Claiming that all English nominals make a paradigm-relevant category distinction on the basis of a few pronouns showing a deviant behaviour means deriving a generalization from an obvious exception to the rule. Such practice is uneconomical with respect to the description of inflectional paradigms, and is, moreover, not particularly likely to make a reasonable approximation to the actual cognitive representation of morphology. On the other hand, the applicability and the lexical scope of the distributional method also depends on typological factors: languages with a high degree of bi-uniqueness in exponency may allow the exceptional presence or absence of a given case marker (associated with a clearly defined range of functions) in the pronoun paradigm to be directly interpreted as a deviation in category oppositions, whereas in languages with a strong tendency for allomorphy and syncretism (both intraparadigmatic and interparadigmatic), a far-reaching distributional analysis remains indispensable.

In sum, there are many languages in which it appears more in keeping with the morphological facts to argue that certain subsets of nominals are sufficiently autonomous as to develop category sets of their own. Such paradigmatic autonomy is only given if the grouping in question is clearly set off from the rest of the nominals as a well-defined natural class of lexemes, based exclusively on shared

[4] Such an analysis has been argued for by, among others, Goddard (1982) and Comrie (1986; 1991). However, cf. also fn. 5 below. The opposite position has been implicitly adopted by Nichols (1992), while Carstairs (1987) refrains from taking an ultimate stance on the issue.

functional or semantic properties (i.e. not on purely language-specific criteria, such as morphological behaviour or phonological shape). Ideally, all lexemes having the relevant functional specifications, and none lacking them, should exhibit the same case-categorial deviation. The personal pronouns are the first and foremost example of such morphology-independently defined natural classes, which will henceforth be referred to as **NP-types** (others are e.g. proper names, or animates vs. inanimates). Paradigm types with more random lexical membership, such as many declension classes in fusional languages, are better accounted for distributionally.

Furthermore, it still needs to be addressed how to identify the language-specific norm against which any deviant category inventory is to be measured. The present analysis takes its cue from Wurzel's (1984) Natural Morphology framework, and particularly from that author's concept of **system-defining structural properties**. In this approach, the inflectional apparatus of a given language at large is characterized by a number of typological cornerstones that receive their normative values from the specific behaviour of the paradigm classes with high type-frequency and unrestricted lexical productivity. Certain paradigm types with lower lexical yield, on the other hand, may well be inconsistent with such a general norm; their evidence has no impact on the outlines of the language-specific morphology as a whole. In this sense, the overall category system of a language operates independently from the quirks of paradigmatically isolated, unproductive NP-types.[5] Now a more fine-grained definition of case-asymmetry can be formulated:

A given language shows case-asymmetry if the inventory of morphological case categories in at least one of its distinct NP-types deviates from the case inventory in its normative system.

For present purposes, case categories are defined as exclusively morphological units that constitute (other things being equal) formally distinctive opposition terms within inflectional paradigms. Of course, such unique formal expressions are necessarily associated with a certain range of syntactic or semantic functions, but it is always distinctiveness in form, rather than function, that gives rise to separate cases. This has repercussions on the question of polysemy vs. homophony in constellations where formally identical case expressions cover a broad range of functions: if such identity of expression is a constitutive feature of the language's normative system, it is interpreted as a single, though multifunctional category. Consequently, it has no impact on this purely morphology-based assessment even if the various component functions of such a comprehensive category differ in their syntactic behaviour (e.g. with a hypothetical combined ergative-function, the

[5] In Wurzel's view, the system-defining structural properties determine the directionality of future language change. This assertion often does not stand the test of empiricism, but is irrelevant to present purposes anyway.

ergative occurrences may control person agreement on the verb while the instrumental occurrences do not).

It should be highlighted once again that case-asymmetry is only concerned with the inflectional potential of nominal lexemes, i.e. with category structure and size of word-specific paradigms. Case-marking asymmetries not affecting the paradigm, such as non-canonical case assignment depending on definiteness, aspect, or clausal polarity fall outside the definitional scope of the concept.

16.3 Subtypes of case-asymmetry

Instances of case-asymmetry can be typologically subclassified according to various parameters, viz. the polarity of the deviation from the normative system (more, less, or simply other categories, and equipollent-quantitative asymmetry),[6] as well as further structurally-based subdivisions, the NP-type in which the asymmetry occurs, and the functional class of the case category in asymmetrical representation (core-argument, peripheral, and adnominal categories). These logically independent parameters form a multidimensional grid. Note that case-asymmetry as such is a property of an entire language (in the sense that the normative systems may or may not have scope over the whole nominal vocabulary), but its logical subtypes are better viewed as local phenomena, given that the same language may exhibit various instances of asymmetry simultaneously. The following overview follows a primarily structure-oriented classification pattern, beginning with straightforward **additive**(-quantitative) asymmetry:

(2) additive case-asymmetry, schematic[7]

	normative system	**personal pronouns**
Case A (default category)	-Ø	-Ø
Case B		-x
(Case C)	-y	-y

[6] Equipollent asymmetry is a subtype that does not become manifest in the pronouns by definition (cf. Iggesen 2005b: 351–4).

[7] In the schematic tables (2), (4), (6), (8), and (11), the variables -x, -y, -a, -b represent distinct case categories, ideally coded by distinct exponents, whereas -Ø stands for the morphologically zero-marked base category, and -xy in table (4) refers to a multifunctional, though unified, case. The dotted separation in table (6) implies there is a unified inflectional category, but two different syntactic constructions depending on the case functions (x# being a free morpheme that may or may not be cognate with the bound case marker used to code case -x in the rightmost column).

978–0–19–920647–6 17-Spencer-ch16 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 251 of 258 May 16, 2008 18:58

This is the pattern instantiated by English and the other European languages mentioned earlier. It is the most common subtype of case-asymmetry cross-linguistically, and is arguably the least prone to controversy with respect to the limited lexical applicability of the category involved. Particularly prominent are constellations with additional core-argument (in general, object-marking) cases, i.e. the immediate parallels to the European prototype. A somewhat less usual illustration of this subtype is the pronominal case opposition in Haida, corresponding to a split-intransitive alignment pattern (i.e. agentive vs. patientive):

(3) Haida (Levine 1977: 172–3)

	'child'	1 SG.	2 SG.	1 PL.	2 PL.	3 SG./PL.
AGT	ğaxa	ła	da	t'alʌŋ	*dalʌŋ*	la
PAT		di	dʌŋ	ʔ/iλ̃)'ʌ	*dalʌŋ*	lʌ

An uncommon manifestation of this subtype is also found in Washo (Jacobsen 1964: 445), inasmuch as here the distinctive objective case is restricted to the third person pronouns (of all numbers) rather than occurring with the SAP persons. An asymmetrically represented genitive exists in the Papuan language Usan (Reesink 1987: 53), where all personal pronouns (but only these) have such case forms marked by *-nou* (e.g. 1 SG. *ye* vs. *yonou*). An example of an additional peripheral case is the comitative in Spanish, limited to three pronominal lexemes (e.g. 1 SG. [*con*]*migo*).

Another recurrent pattern of additive case-asymmetry is **split-function asymmetry**. In such constellations a multifunctional case category of the normative system corresponds to (at least) two distinct categories in the deviant NP-type. Conversely, the sum of the functions expressed by the separate cases in question is covered by an overarching category in the productive classes of full nouns. By stipulation, this multifunctional category is not the unmarked base case of the paradigm.[8]

(4) split-function additive case-asymmetry, schematic

	normative system	**personal pronouns**
Case A (default category)	-Ø	-Ø
Case B	-xy	-x
Case C		-y

Instances of split-function asymmetry are often considered case-syncretic (within the normative system) (cf. Araona in Baerman's chapter in this volume). The

[8] Otherwise, there would be no formal distinction between this subtype and straightforward additive case-asymmetry, as represented schematically in table (2).

methodology applied here, however, entails that the differentiation of forms found in a minority NP-type cannot be taken as basis for identifying the category inventory of the normative system. A concrete example of split-function asymmetry is the pronominal declension in Armenian, involving a split between the otherwise identically coded dative and genitive:

(5) Armenian (Kozintseva 1995: 13), partial paradigm

	'wolf'	1 **SG.**	2 **SG.**
NOM	gayl	es	du
DAT	gayli	inj	k'o
GEN		im	k'ez

Split-function constellations occasionally break up also rather well-defined single case of the normative system into separate expressions of more fine-grained component functions. This happens in Lithuanian, where the full nouns have a single genitive, but certain pronouns distinguish an adnominal genitive from a genitive of complements (Mathiassen 1996: 72).[9]

A subtype with arguably much less impact on the overall morphosyntactic system is **case-asymmetry with functional equivalent**. Here the relational function in asymmetrical representation evinces unique coding for all nominal lexemes. However, only in the deviant NP-type this coding qualifies as inflectional case, whereas with all other nominals, non-bound markers (adpositions) or other devices outside the definitional scope of nominal case are employed instead.

(6) additive case-asymmetry with functional equivalent, schematic

	normative system	**personal pronouns**
Case A (default category)	-Ø	-Ø
Case B	-Ø# x# (no separate case status)	-x
Case C	-y	-y

An example in point is the Papuan language Una: here the ergative-instrumental function is expressed in the normative system by means of the postpositions *aji* (with common nouns) and *beji* (with proper names), whereas in the personal pronouns fusional forms are used (e.g. 1 SG. *ni* ABS vs. *niji* ERG) (Louwerse 1988: 108–9).

[9] Admittedly, the Lithuanian genitive appears functionally well-defined only from the Indo-European perspective.

(7) Una (Louwerse 1988: 108–9), partial paradigm

	male name	1 SG.	2 SG.	3 SG.
ABS	Endok	ni	kan	er
ERG(/INST)	(Endok beji)	niji	kanci	erci

The mirror image of additive asymmetry is **subtractive**(-quantitative) case-asymmetry. Here the paradigm of the deviant NP-type contains fewer case categories than the normative system. The abstract pattern can be visualized as follows:

(8) subtractive case-asymmetry, schematic

	normative system	**personal pronouns**
Case A (default category)	-Ø	-Ø
Case B	-x	
Case C	-y	-y

The subtractive subtype may seem to contradict the general methodology: the exceptional absence of an overt case-form distinction is particularly susceptible to an interpretation as homophonous expressions of distinct underlying categories (i.e. as case-syncretism), given that additional arguments are necessary to make a point that the case inventory of the general system is not applicable to the NP-type in question. Iggesen (2005b: 457–95) illustrates that there are indeed many instances which inherently allow for both analyses, albeit with various degrees of plausibility (hence, the concrete case-marking constellations build a continuum between asymmetry and syncretism). There are, however, also situations where an analysis as syncretism appears inappropriate or artificial, such as the exceptional lack of an otherwise bi-unique case marker, which is replaced for the function in question by the unmarked default form. In view of the overall typology of such languages, it may be quite far from structural reality to posit a zero allomorph of an otherwise distinctive case morpheme (the former considered only incidentally homophonous to the basic case category). It is much more in keeping with such language structures to derive the non-existence of the respective category in the minority paradigm directly from the non-applicability of its single marker.[10]

[10] Baerman (in this volume) discusses the interesting example of Luiseño, where inanimate nouns are not marked for accusative, while their adjectival modifiers are; thus, the case function remains differentiated syntagmatically. This may be interpreted as a mismatch between case as morphological (at the word level, concerning the noun) and as syntactic (at the phrase level) operation. However, further difficulties can then arise e.g. for the assessment of German, where the case distinctions are chiefly coded on the phrasal determiner.

978–0–19–920647–6 17-Spencer-ch16 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 254 of 258 May 16, 2008 18:58

The clearest examples in point are found among semantically defined subclasses of the full nouns rather than in the personal pronouns, given that the latter have a greater tendency for structural irregularity in their paradigmatic base forms.[11] With respect to the pronouns, the Chittagong dialect of Bengali can be cited, a language with only a low degree of allomorphy in its case markers. One of these is the ergative suffix -(*j*)*e*, whose occurrence is restricted to past-tense clauses (following tri-partite alignment). The personal pronouns, however, do not inflect for the ergative, even though full nouns having similar phonological shapes in their base forms do.

(9) Chittagong Bengali (Učida 1970: 31, 38–9)

	'brother'	**1st sg.**	**1st pl.**
NOM	baɪ	ãɪ	ãra
ERG	baije		
OBJ	baijɔre	ãre	ãrare

Case-asymmetry with functional equivalent takes on a more significant role in the realm of subtractive polarity, in view of the fact that the replacement of an inflectional case by some equivalent non-case construction cannot be mistaken for syncretism. Subtractive asymmetry with functional equivalent often affects the genitive. In many languages this case is substituted for in the pronominal domain either by possessive adjectives (defined as agreement-bearing forms) or personal head-marking on the possessee, neither of which qualifies as genitive form of the free pronouns. The former substitution strategy can be illustrated by Murle (Nilo-Saharan), where pronouns in possessor function occur in two sets of forms according to the gender of the possessee.[12]

[11] There are languages in which the well-delimited NP-type of the toponyms is excluded from inflection for the stative-locational case, e.g. the Papuan language Yimas (Foley 1991: 170–1).

[12] Adjectival replacement of the genitive in possessor function is not necessarily tantamount to case-asymmetry, though. In the Slavic languages, pronominal possessors are expressed by agreement-bearing personal adjectives. These items constitute independent lexemes, and typically have full-fledged case paradigms of their own. However, the Slavic genitive extends to other syntactic functions, such as the complements of certain prepositions (cf. also Lithuanian above). In such non-possessive functions the personal pronouns show genuine case forms, which represent the genitive within the paradigm. Consequently, the overall case inventory of this NP-type remains the same as in the normative system.

(10) Murle (Arensen 1982: 58, 93–4, 98)[13]

	'crocodile'	1 **SG.**	2 **SG.**
OBJ	agul	aneeta	ineeta
NOM	aguli	naana	niina
GEN	agulo	cun ∼ onan (sg.) cigan ∼ ogan (pl.) (no case status)	cun ∼ onun (sg.) cigun ∼ ogun (pl.) (no case status)

There are also case-asymmetry constellations in which the NP-type representing the asymmetry domain differs from the normative system not by having a deviant number of case categories, but rather by simply having a different set of categories, i.e. categories that cannot be interpreted as identical across the nominal lexicon for both functional and morphological reasons. This means that the mapping relationship between (certain) distinct forms and the range of occurrences covered by them is substantially dissimilar between the two paradigm sets, whereas the count of categories typically gives the same result. This subtype is called **qualitative** case-asymmetry. It can be modelled as follows:

(11) qualitative case-asymmetry, schematic

	normative system	**personal pronouns**
case function 1	-x	-a
case function 2		-b
case function 3	-y	

Cross-linguistically, qualitative asymmetry is virtually always realized in form of alignment splits running neatly along the boundaries of personal pronouns and other nominals (e.g. 'split-ergativity'). This coding situation is common in Australian languages, where it may be seen as an areal feature. Often interpreted as case-syncretic following Goddard (1982), these languages clearly meet the definitional criteria of case-asymmetry established above. An illustrative example is Pitjantjatjara:

[13] In this table, the shaded background indicates that the forms given in these cells do not belong to the case paradigm of the nominal lexemes in question.

[14] In this table, the leftmost column refers to syntactic relations, while the columns immediately to the left of the two word-form columns indicate the paradigmatic case categories (as can also be seen from the different category labels used for the two subsystems). Categories and forms involved in the asymmetry constellation are additionally italicized.

978–0–19–920647–6 17-Spencer-ch16 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 256 of 258 May 16, 2008 18:58

(12) Pitjantjatjara (Bowe 1990: 9–10, 12)[14]

		'man'		1 SG.
A relation	*ERG*	*watingku*	*NOM*	*ngayulu*
S relation	*ABS*	*wati*		
O relation			*ACC*	*ngayunya*
possessive funct.	GEN	watiku	GEN	ngayuku

16.4 Motivations

Case-asymmetry is a relatively abstract property, it occurs with opposed polarity values, and affects functionally diverse sets of cases. All this indicates that a single, overarching principle explaining all instances is not likely to exist, but rather that the individual constellation types have separate (though partly interrelated) motivations. Likewise, competing explanations may often be seen as reinforcing each other rather than as mutually exclusive. The most significant motivation structures contributing to the emergence of case-asymmetry are economy (frequency and efficiency), pragmatic factors, and the special functions of personal pronouns. Furthermore, many instances of case-asymmetry are correlated with the universal nominal hierarchy (Silverstein 1976).

The fact that case-asymmetry is so often found in SAP pronouns, but rarely in other NP-types, is to be explained by the functional separateness of this word class from other nominals. Personal pronouns denote ever-shifting speech roles in dialogue (Bhat 2004: 6). Their referentiality is situationally concrete, but of a very different kind from the anaphoric or deictic coreference to established entities of third-person pronouns and demonstratives. Reference to the individual speaker / hearer (as entities) is not the primary task of the SAP pronouns. In this sense, they are ultimately pragmatic and grammatical function words, and may therefore be viewed as the least prototypical nominals. Hence, if a language develops case-asymmetry at all, there is rightfully a strong tendency to do it precisely in this functionally most autonomous NP-type. Third-person pronouns are somewhat closer to full nouns through the link of potential coreference, and therefore tend to pattern like these in their category inventories.

The SAP pronouns also occupy the highest positions in the nominal hierarchy, which is itself based on a prototype-oriented cluster of animacy, empathy,

topicality, etc. In addition, they tend to be among the most frequent nominals in discourse.[15]

The relevance of the nominal hierarchy for **differential object marking** (Bossong 1985c) has often been observed. Two major constellations of pronominal case-asymmetry are indeed special manifestations of differential object marking: additive asymmetry affecting object cases, and qualitative asymmetry. Here relative unexpectedness is the most likely rationale: the deviant marking is reserved for nominals that are less prototypically found to act as semantic patients / O arguments, and where the hearer would be therefore most likely to misinterpret the speaker's utterance intention if not guided otherwise. The relative frequency of pronouns ensures that differential marking is not a waste of resources; on the contrary, it is maximally economical in the sense of a purpose-bound distribution of structural 'expenses'. Subtractive asymmetry, on the other hand, runs against the nominal hierarchy, and so does the object case restricted to third-person pronouns in Washo. Additive asymmetry involving human-oriented non-object cases in the pronouns may also be viewed as correlated with the hierarchy, though in the sense of a semantic match between case relation and the most appropriate class of referents (i.e., prototypicality) rather than via unexpectedness (Iggesen 2005b: 606–7). Such human-oriented categories are e.g. genitive, comitative, or dative.

Asymmetry with functional equivalent is best accounted for by frequency arguments and prototypical matching. This holds equally true for subtractive asymmetry with a genitive equivalent: humans are the most typical (and most frequent) possessors of entities; hence, the very tight integration of their pronominal representation with the possessee in form of head-marking morphology (through affixation) or their high degree of syntagmatic indexicality (through adjectival agreement) can also be seen as semiotically advantageous.

[15] The validity of this impressionistic statement varies with language types (non-pro-drop vs. pro-drop), the individual person–number combinations (first person sg. pronouns are certainly more token-frequent than second person plural), and discourse genres.

PART III

SYNTAX OF CASE

CHAPTER 17

CASE, GRAMMATICAL RELATIONS, AND SEMANTIC ROLES

BEATRICE PRIMUS

This chapter presents an overview of the basic issues concerning the relationship between cases, grammatical relations, and semantic roles such as agent and patient.[1] Section 17.1 reviews general assumptions about this relationship. In most approaches, semantic roles are directly linked to abstract grammatical relations for the core arguments of the clause. Cases are considered to be a surface expression of grammatical relations. Accordingly, this article deals primarily with the relationship (or linking) between grammatical relations and semantic roles in different types of approaches (sections 17.2–17.4). The particular contribution of cases is mentioned at the end of sections 17.3 and 17.4. Section 17.5 summarizes the main results of this overview.

[1] Besides semantic roles and closely related event-structural notions such as change of state, there are other semantic factors that may determine the choice of grammatical relations. These factors include, among others, topic, animacy, (in)definiteness, and negation (e.g. Hopper and Thompson 1980, de Hoop 1996, Kittilä, this volume).

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 262 of 275 May 16, 2008 18:58

17.1 Transparency principles

Approaches dealing with this topic incorporate principles that guarantee a transparent relationship between semantic roles and grammatical relations, but may differ with respect to how strictly they adhere to transparency. The strict principles in (1)–(2) characterize the framework of mainstream generative grammar:[2]

(1) The Theta-Criterion: Each argument bears one and only one theta-role, and each theta-role is assigned to one and only one argument (Chomsky 1981: 36).[3]

(2) The Uniformity of Theta-Assignment Hypothesis (UTAH): Identical thematic relationships between items are represented by identical deep structural relationships between those items (Baker 1988: 46).[4]

Within functional-typological research, transparency is ensured by two violable general functions of case marking, the distinguishing (or discriminatory) and the characterizing function (e.g. Comrie 1981a: 117–20, Dik 1997: 369–70, Song 2001: 156–67). In their distinguishing function, cases are used for discriminating concomitant constituents with different semantic or syntactic properties. This function is claimed to explain the fact that the subject of a one-place predicate is generally marked by the same case, the nominative or absolutive, irrespective of its semantic role. Ideally, a second case is only used with predicates selecting at least two roles, as in Japanese and Dyirbal, and a third case is only required by predicates with at least three roles.

The second broad function of cases is the characterizing one. Ideally, a case has this function if it is used for all and only the noun phrases with a certain type of semantic role or syntactic function. The syntactic characterizing function is a common assumption in traditional grammars. They define subject, direct object, and indirect object in terms of the nominative, accusative, and dative respectively (cf. Maling 2001 for arguments against this view). As to the semantic characterizing function, argument realization is universally restricted by semantic roles within certain limits, as we shall see in the following sections.

Due to the characterizing function, case selection may be more variable than predicted by the distinguishing function. Thus, for example, the case selected by one-place predicates may vary according to the semantic role of the argument

[2] Similar principles are the Universal Alignment Hypothesis of Relational Grammar (Perlmutter and Postal 1984) and the Function-Argument-Biuniqueness Condition of Lexical Functional Grammar (cf. Butt 2006: 128). See Jackendoff (1990), Dowty (1991), and Culicover and Jackendoff (2005) for arguments against such strict one-to-one mapping assumptions.

[3] In the Minimalist Program (cf. Chomsky 1995, Chap. 4), the Theta-Criterion is derived from the principles that constrain feature checking.

[4] Later, Baker (1997: 104) is more cautious. He refers to similar thematic roles in similar initial structural relations and to Dowty's (1991) approach (cf. section 17.3 below) without adopting it in a strict sense.

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 263 of 275 May 16, 2008 18:58

(e.g. German *mir war kalt* 'I (DAT) felt cold' vs. *ich arbeitete* 'I (NOM) worked'). But the variation induced by the characterizing function is usually constrained by the distinguishing function. Let us consider the dative as the third case in addition to the nominative and accusative in Latin (cf. Dik 1997: 370) and German (cf. Primus 2006b). Very few one-place predicates select the dative instead of the nominative (less than 1 per cent). The number of two-place predicates with a dative in addition to a nominative increases to *c.* 7 per cent, while the ratio of three-place verbs selecting a dative reaches *c.* 70 per cent and all these verbs also select an accusative and a nominative.

Transparency principles and functions presuppose more specific assumptions about semantic roles and their linking to cases or grammatical relations. There are several views on semantic roles. In one view they are unanalysable entities that are listed in the lexical representation of a predicate (cf. section 17.2). The other view is decompositional: a few general roles are defined in terms of basic notions such as causation, motion, or sentience (cf. section 17.3) or in terms of structural positions in the lexical representation of verbs (cf. section 17.4).

17.2 Grammatical relations, role lists, and role hierarchies

Many influential approaches use a list of informally defined semantic roles and a linking mechanism that is grounded on a role hierarchy. Fillmore's Case Grammar (1968, 1977) is a pioneering work in this line of research. The semantic roles, 'deep cases' in his terminology, are characterized relative to the action or state of affairs identified by the verb as follows (1968: 24–5): Agentive for the typically animate instigator; instrumental for the inanimate force or object causally involved; objective as the semantically most neutral role; dative for the typically animate being affected by the state or action; factitive for the object or being which results from the action or state. The list of deep cases, which also includes locative, is considered preliminary, but Fillmore is confident that they form a small set (1968: 5). This last assumption is needed in this kind of approach for the sake of one-to-one mapping principles.

The lexical entries of verbs indicate their selection of roles, as shown in (3). Parentheses indicate optional roles (1968: 27, 35):

(3) *open:* [__ objective (instrumental) (agentive)]
show, give: [__ objective + dative + agentive]

The basic structure of sentences includes deep cases,[5] which are linearly ordered as in the lexical entries of verbs. Surface structures, in which one deep case is realized as a syntactic subject, are derived according to the following rule for the unmarked subject choice (Fillmore 1968: 35):

(4) If there is an agentive, it becomes the subject (cf. 5a,b);
otherwise, if there is an instrumental, it becomes the subject (cf. 5c);
otherwise the subject is the objective (cf. 5d).

In accordance with this rule, the verb *open* can be used in the following surface structures:

(5) a. *John opened the door with a key.*
b. *John opened the door.*
c. *The key opened the door.*
d. *The door opened.*

In order to explain the role asymmetries needed for subject and object selection, many role-list approaches use a role hierarchy. The hierarchy that motivates Fillmore's subject selection rule is (6):

(6) agent > instrumental > objective

This kind of treatment of semantic roles and linking characterizes many influential approaches including Dik's Functional Grammar, Bresnan's Lexical Functional Grammar, and Grimshaw's approach within mainstream generative grammar.[6] Their role hierarchies are shown in (7), (8), and (9) respectively:

(7) agent > goal (i.e. patient or theme in other approaches) > recipient > beneficiary > instrument > location > time (Dik 1997: 37)

(8) agent > beneficiary > recipient / experiencer > … > patient / theme > location (Bresnan 2001: 11)

[5] Some recent approaches also view semantic roles as syntactic notions (e.g. Hale and Keyser 1992 and most contributions in Erteschik-Shir and Rapoport 2005). Arguments against this view are presented in Butt (2006: 31, 49–51).

[6] Parsons' Neo-Davidsonian approach (1990, 1995) also uses a role list. Davidson (1967) and Parsons are concerned with the proper representation of events. A sentence like *Brutus stabbed Caesar in the agora* designates an event and entails the following: *There was a stabbing; The stabbing was by Brutus; The stabbing was of Caesar; The stabbing was in the agora*. In order to account for such entailments Parsons represents semantic roles as relations between events and things (or individuals) and connects the event components by logical conjunction. Cf. (a), where each conjunct represents the above-mentioned entailments of *Brutus stabbed Caesar in the agora* (cf. Parsons 1995: 636):

(a) $(\exists e)[\text{Stabbing}(e) \,\&\, \text{Agent}(e, \text{Brutus}) \,\&\, \text{Theme}(e, \text{Caesar}) \,\&\, \text{In}_{\text{Location}}(e, \text{the agora})]$

As Parsons admits (1995: 639–40), his treatment does not offer any substantive information on semantic roles and argument realization. Nevertheless, Neo-Davidsonian representations are still used, particularly by those concerned with event structure and its grammatical consequences (e.g. Krifka 1989, 2004; Engelberg 2000).

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 265 of 275 May 16, 2008 18:58

(9) Agent (Experiencer (Goal / Source / Location (Theme)))
(Grimshaw 1990: 8)

In Functional Grammar and Lexical Functional Grammar, the role hierarchy is mapped onto abstract, non-structural syntactic functions. It is also used to determine the basic order of verbal arguments (e.g. Siewierska 1988, Dik 1997). Grimshaw uses a structured role hierarchy that is transparently projected onto structural syntactic functions if event-related factors such as change of state do not interfere.[7]

The logic of the subject selection principle in these approaches is the same as Fillmore's in (4) above: if the verb selects the first role in the hierarchy (i.e. an agent), this role becomes the subject; otherwise, if the verb selects a second role in the hierarchy, this role becomes the subject, and so forth rightwards in the role hierarchy. Object selection is based on the same hierarchy and selects the patient or theme as the first choice.

Role lists are also used in mainstream generative grammar (cf. Chomsky 1981: 34–54, 1995: 30–3, 186–91; Baker 1988, 1997). Baker (1997: 120–1) proposes the following universal mapping principles that are meant to specify the UTAH in (2) above:

(10) a. An agent is the specifier of a higher VP.
b. A theme is the specifier of the lower VP.
c. A goal, path, or location is the complement of the lower VP.

The structural analysis in (11) illustrates Baker's assumptions (cf. section 17.4 for stacked VPs):

(11) $_{VP}$[*John* $_{VP}$[*the ring* $_{V'}$[*passed to Mary*]]]

Baker's mapping principles in (10) are meant to eliminate a role hierarchy as a basic concept. They guarantee that agents are higher in the structural hierarchy than themes, and themes higher than goals, paths, or locations, as shown in (11). An advantage of structural mapping principles is the specification of the absolute, i.e. exact structural position of an argument.

Let us summarize. Role lists are a convenient tool for preliminary role analyses and linking hypotheses. Their main weakness is that roles are not defined in terms of a limited, conceptually well-motivated set of semantic primitives. There are three negative consequences of this weakness. First, the number of individual roles exceeds by far the number of core syntactic functions, so that one-to-one transparency principles cannot be taken literally. Consider for example the alternative roles for one hierarchy slot or syntactic position in (8), (9), and (10) above. If one includes semantically underspecified predicates (cf. (12) and (13) below), the number of alternative roles for one position expands considerably. Role-list approaches are forced to claim that many alternative roles belong together in some way that is left unexplained.

[7] Event-related factors are introduced for causative-inchoative psychological verbs such as *frighten* and *please* (Grimshaw 1990, Chap. 2).

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 266 of 275 May 16, 2008 18:58

This leads us to the second problem. Role lists cannot offer an understanding of whether two roles count as identical for linking or not. For the sake of transparency, such approaches underestimate role differences and this may lead to empirically inaccurate analyses. Let us consider Baker's assumptions about ditransitive verbs. He collapses possessive and locative roles (cf. (10c) above) and derives the possessive construction (e.g. *John passed Mary the ring*) by syntactic movement from the locative construction (cf. (11) above). However, approaches focusing on the meaning of such predicates have established that verbs such as *pass* or *give* have two readings: a possessive reading with a recipient (or abstract goal) and a locative reading with a locative goal (cf. Wunderlich 1997, Krifka 2004, Levin and Rappaport Hovav 2005: 206–28). The two meanings are not equivalent, but related: the possessive construction entails the locative one. The entailment explains why a fairly large number of verbs occurs in both constructions, but the truth-conditional meaning differences cast a doubt on Baker's assumption that they are mere suggestions (1997: 89).

The third problem of this kind of approach is that role lists are unstructured sets, so that additional stipulations in form of role hierarchies or mapping principles yielding role asymmetries are needed. However, these additional stipulations do not offer a substantive explanation for the attested asymmetries. As a result, role hierarchies and mapping asymmetries come in an unduly high number of variants, as shown above (cf. Levin and Rappaport Hovav 2005: 154–83, Primus 2006a). As we shall see in the following sections, there are decompositional semantic approaches that explain role asymmetries on semantic grounds and reduce the inventory of superordinate role concepts dramatically without neglecting finer distinctions.

17.3 Grammatical relations and proto-roles

In this section we will discuss Dowty's (1991) work as a representative of a non-structural decompositional approach. Related influential approaches are Role and Reference Grammar (cf. Van Valin and LaPolla 1997; Van Valin, this volume) and the transitivity concept of Hopper and Thompson (1980; Kittilä, this volume). Dowty defines two superordinate proto-roles by a small set of semantic primitive properties (Dowty 1991: 571–2). The agent proto-role is characterized as follows:

(12) Agent proto-role:
   a. x does a volitional act: *John refrains from smoking*.
   b. x is sentient of or perceives another participant: *John knows / sees / fears Mary*.

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 267 of 275 May 16, 2008 18:58

c. x causes an event or change of state in another participant: *His loneliness causes his unhappiness.*
d. x is moving: *Water filled the boat.*
e. x exists independently of the event named by the predicate: *John needs a car.*

Although most verbs select more than one proto-agent property for their subject argument (e.g. *murder*, *nominate*, or *give*), each of these properties can occur in isolation as shown by the subject argument in the examples in (12a)–(12e). The patient proto-role is defined and illustrated by the object argument of the examples in (13):

(13) Patient proto-role:
a. x undergoes a change of state: *John moved the rock.*
b. x is an incremental theme: *John filled the glass with water* (also stationary relative to other participants).[8]
c. x is causally affected by another participant: *Smoking causes cancer.*
d. x is stationary relative to another participant: *The bullet entered the target.*
e. x does not exist independently of the event, or not at all: *John needs a car/ seeks a unicorn.*

The list of properties in (12) and (13) is preliminary for Dowty: properties can be deleted or added without changing the logic of his approach. Candidates for deletion are stationary (cf. Primus 1999: 42–3) and incremental theme (cf. Levin and Rappaport Hovav 2005: 106–10). Additions include the target of sentience or emotion (cf. Pesetsky 1995), possessor (proto-agent) and possessed object (proto-patient) following, among others, Jackendoff (1990) and Wunderlich (1997). These changes lead to the following simpler and more consistent system (cf. Primus 1999, 2006a). The two proto-roles involve the same concepts in converse pairs: volitionally acting or causing vs. volitionally or causally affected, moving (i.e. physically active) vs. moved (i.e. physically changed), sentient vs. target of sentience, possessor vs. possessed, independent vs. dependent existence.

The specific roles of role-list approaches can be defined in terms of proto-role properties: agents by volition and possibly more proto-agent properties; instruments and causers by causation without volition; experiencers by sentience without other properties. To sum up, decomposition allows for a high number of specific roles to be subsumed under a small set of general roles, thus combining differentiation with generalization in an elegant way.

[8] Incremental theme is an event-related role introduced by Krifka (1989) for a participant whose degree of affectedness parallels the degree of completeness of the event. Incremental affectedness does not imply a physical change of state, cf. *read a book* and *memorize a poem*. Conversely, not every change of state or location is incremental, cf. *push a cart*.

Syntactic argument realization is assumed to be sensitive to the higher or lower number of semantic basic properties accumulated by an argument. Dowty's argument selection principle is stated as follows (1991: 576):

(14) The argument for which the predicate entails the greatest number of proto-agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of proto-patient entailments will be lexicalized as the direct object.

The principle is meant to capture lexical default mappings for arguments with a high number of consistent properties such as selected by the verbs *break* and *hit*. Underspecified roles that accumulate a low number of consistent proto-role properties or none at all may have a variable realization, as will be shown later in this section.

Principle (14) holds only for accusative languages, which include most European and African languages. For ergative languages, such as Basque, the Caucasian languages, most Indo-Iranian, Australian, and Polynesian languages, the syntactic association is reversed: the argument with the greatest number of proto-agent properties bears the more marked ergative case; the argument having the greatest number of proto-patient entailments is in the least marked case, i.e. the absolutive or nominative (cf. Dowty 1991: 582). This account explains the well-documented fact that the ergative–accusative distinction is most clear-cut with transitive verbs that select a high number of consistent proto-role properties (e.g. Hopper and Thompson 1980; Primus 1999, Chap. 4; Malchukov 2005).

The accusative pattern is amply illustrated by the English examples in this chapter and can also be captured by the linking rules of role-list approaches (cf. (4) and (10) above). (15) shows the ergative case pattern in Samoan (Mosel 1987: 455):

(15) *Sa fasi le tama e le fafine*
PAST hit DET boy ERG DET woman(ABS)
'The woman hit the boy.'

In the ergative construction, the agent is expressed by a more marked case, i.e. the ergative, while the patient is coded by the least marked case, i.e. the nominative or absolutive. The proto-role approach predicts that this construction is most consistently used with roles that accumulate a high number of consistent proto-role properties. Such roles are selected by verbs such as *break* and *hit*.

The typological ergative–accusative distinction is explained by the logic of Dowty's argument selection mechanism, which yields a reversible role hierarchy (cf. Primus 1999, 2006a). In other words: agents and patients are on an equal footing if the high number of consistent properties is the main determinant of subject and object selection. Cf. the two inverse hierarchies in (16a,b), in which a role on the left is more accessible to the case in question than its follower to the right:

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 269 of 275 May 16, 2008 18:58

(16) a. Nominative or ergative: agent > experiencer / possessor > recipient-like roles > target of sentience / possessed object > patient
b. Accusative or absolutive: patient > possessed object / target of sentience > recipient-like roles > experiencer / possessor > agent

By contrast, role-list approaches do not offer an explanation for the inverse hierarchy in ergative languages. These approaches collapse the two hierarchies that are needed for subject and object selection and for the ergative–accusative distinction into one unidirectional hierarchy.

Intransitive predicates generally use the nominative or absolutive for all types of roles, as predicted by the distinguishing function of cases (cf. section 17.1). This function competes with Dowty's principle, which can be viewed as a specification of the characterizing function. It predicts role-driven case variation in intransitive clauses. This pattern, called *active* or *split intransitive,* is illustrated by the following Guarani examples (Gregores and Suárez 1967: 110):

(17) a. *a-ma.apo*
1SG.A-work
'I work.'
b. *še-manuʔa*
1SG.P-remember
'I remember.'

A- vs. P-markers are selected depending on the higher vs. lower number of agentive properties of the argument (cf. Primus 1999, Chap. 4). Nearly 100 per cent of the volitional agents are coded by an A-marker. This marker is shown in (17a). 89 per cent of the fully underspecified roles are indicated by a P-marker. This marker is illustrated in (17b). Fully underspecified roles do not have agentive properties besides independent existence. They are selected by such verbs as *porã* 'be beautiful, right' and *marete* 'be powerful, strong'.

A proto-role approach captures the well-attested fact that arguments with a small number of consistent role properties exhibit a variable realization cross-linguistically or within one language (cf. Malchukov 2005; the contributions in Aikhenvald et al. 2001; Bhaskararao and Subbarao 2004). The following Japanese example (Shibatani 2001: 319) and its English translation illustrate the coding variation for psychological predicates:

(18) *Sensei ni eigo ga wakaru*
teacher DAT English NOM understands
'The teacher understands English.'

In the Japanese example, the experiencer is expressed by the dative, while the target of sentience is in the nominative. In the English translation, the experiencer is in nominative subject position and the target in the accusative object position.

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 270 of 275 May 16, 2008 18:58

This kind of variation is typical for non-causative psychological predicates such as *understand*, *love*, and *fear*, which select an experiencer and a target of sentience (cf. Pesetsky 1995, Levin and Rappaport Hovav 2005: 22–3). We limit the following discussion to this type of psychological predicate.

Their coding variation, while corroborating Dowty's view, poses a problem for role-list approaches. The hierarchy agent > patient / target of sentience > experiencer leads to the cross-linguistically inaccurate assumption that targets of sentience surface as nominative subjects. This pattern is illustrated by the Japanese example in (18) above. However, its English translation refutes this linking assumption as a cross-linguistically valid generalization. The alternative hierarchy agent > experiencer > patient / target of sentience accounts for the subject experiencer of the English example (cf. Grimshaw 1990) but fails to capture the Japanese example. The most common solution to this problem is to introduce a distinction between structural and lexical cases, to confine transparency principles to structural cases and to treat oblique experiencers as 'quirky' subjects bearing lexical, idiosyncratic cases.

However, this case distinction blurs a common structural property of experiencers with non-causative verbs that can be explained by their classification as proto-agents. Despite their variable case coding, they are always (or preferably) structural subjects, i.e. they occur sentence-initially in terms of basic order as shown in the Japanese example. This structural function explains other syntactic subject properties of oblique experiencers in many languages (cf. the contributions in Verma and Mohanan 1991, Aikhenvald et al. 2001, Bhaskararao and Subbarao 2004 as well as Primus 2006a). In Japanese (cf. Shibatani 2001: 317–21), these subject properties include the selection of the subject honorific form and the antecedent function for the reflexive pronoun *zibun*.

The typological ergative vs. nominative variation in the coding of agents poses a similar problem: it exclusively favours case marking and case-related properties such as verb agreement. In both types of languages, agents show a strong tendency to be realized as structural subjects, e.g. to occur before patients in the basic order and to antecede reflexive pronouns (cf. Primus 1999, 2006a). As to this last property, no language has been found in which a proto-patient exclusively serves as an antecedent for a proto-agent.

In sum, proto-agents, e.g. agents and experiencers, show a strong tendency to be realized as structural subjects irrespective of the number of agentive properties they accumulate and irrespective of their case marking. Dowty's approach explains case marking, which is sensitive to the degree of involvement of a participant, better than structural coding, which seems to ignore this semantic distinction. The next section will discuss structural decompositional approaches which offer a better tool for the structural generalizations missed by Dowty.

17.4 Grammatical relations and lexical semantic structures

In this section, we present Wunderlich's (1997, 2006) Lexical Decomposition Grammar, which is based on Kiparsky (1987) and Bierwisch (1988). This choice is motivated by two aspects of this framework. First, it is purely structural. Second, it uses CAUSE and BECOME as basic operators, which are given an explicit truth-conditional interpretation by Dowty (1979) and are also found with minor modifications in many other approaches (e.g. Van Valin and Lapolla 1997, Van Valin, this volume, Levin and Rappaport Hovav 1995, Primus 1999, Culicover and Jackendoff 2005).[9] We will focus on Wunderlich's treatment of ditransitive verbs. See (19):

(19) a. *Er gab ihr einen Apfel.*
he gave her(DAT) an(ACC) apple(ACC)
'He gave her an apple.'
b. *Er zeigte ihr den Dom.*
he showed her(DAT) the(ACC) cathedral(ACC)
'He showed her the cathedral.'

Wunderlich's (1997: 38, 44) lexical representations of the verbs in (19) are shown in (20):

(20) a. λzλyλx [CAUSE(x, BECOME(POSS(y,z)))]
b. λzλyλx [CAUSE(x, (SEE(y,z)))]

Following Bierwisch (1988), Wunderlich splits the lexical representation of a predicate into two parts. The first part is the theta-structure; it contains the lambda-bound variables,[10] which are associated with theta-roles. Only this part of the representation undergoes the linking procedure. The second part contains semantic roles and their characterization in terms of primitive predicates such as CAUSE, BECOME, and POSS. This part is related to the level of conceptual structure (in square brackets for convenience).

The relative prominence of theta-roles in the lambda-structure plays the crucial part in linking. The prominence hierarchy for the ditransitive verbs shown in (20a,b) is 'x > y > z'. The structural dative has the most specific condition: it is linked to a role for which there is a higher role as well as a lower role [+hr, +lr]. Only role *y* satisfies both requirements. The structural accusative is linked to a role that

[9] Overviews are offered by Dowty (1979, Chap. 2) and Levin and Rappaport Hovav (2005: 69–75).

[10] Lambda operators are a formal means to represent the free arguments of a predicate. The order of the variables of a multi-place predicate is reversed in the lambda-structure in order to guarantee the appropriate argument composition in syntax. The lambda-bound event (or situation) variable *s* in Wunderlich's representations is omitted for convenience as it is not pertinent to our discussion.

is dominated by a higher role [+hr]. Two roles, *y* and *z*, are [+hr]. The competition between these roles is decided by a specificity condition: the dative, which is the most specific case matching both features of role *y*, is selected for this role. This leaves role *z* for the accusative. Role *x* is linked to the structural nominative, which is a default case lacking prominence restrictions.

To sum up, Wunderlich's linking principles predict the following general patterns for structural cases: one-place verbs select a structural nominative, two- and three-place verbs select a structural accusative for the lowest argument in addition to the nominative for the highest argument, and three-place verbs select a structural dative for the intermediary argument in addition to the nominative and accusative. These are indeed the default options in accusative languages (cf. Stiebels 2002 for ergative languages in this framework). Cases that do not conform to this structural pattern are assumed to be lexical, i.e. idiosyncratic cases. We will discuss them later in this section.

The abstract structural cases may be expressed according to Wunderlich by morphological cases, as in the German examples (19a,b), or by structural positions, as in English. (21) shows the corresponding structural positions for the three structural cases of the verb *give* (cf. Wunderlich 1997: 39):

(21) $_{VP}$[x CAUSE $_{VP}$[y [POSS z]]]

The stacked VPs, in which each theta-marking basic predicate is a verbal head in syntax, accommodate the assumption of the Minimalist Program that each thematic role is assigned by a specific verbal head within its maximal projection (cf. Chomsky 1995: 172). Thus, structural decompositional approaches offer a substantive semantic motivation for stacked VPs that is lacking in purely syntactic approaches (e.g. Larson 1988) and in role-list approaches (e.g. Baker 1997).

Wunderlich's semantic view on structural cases contrasts with the assumption of mainstream generative grammar that structural cases are dissociated from theta-marking (cf. Chomsky 1981: 171). Closer to the spirit of the generative enterprise is Wunderlich's assumption that structural cases are associated with theta-roles that are defined in purely structural terms. This structural approach is a good means to classify different specific roles by the relational features [±hr] and [±lr], a case at hand being [+hr, +lr] for recipients and addressees (cf. (19) and (20) above). However, the confinement to the structural relations of the lambda-structure leaves some facts unexplained (cf. also Butt 2006: 111–17).

One fact is the split intransitive pattern (cf. section 17.3 above). Recall that in Guarani, an agent regularly takes one type of marker, while a semantically underspecified argument regularly selects the other type of marker. This semantic generalization cannot be accounted for in a straightforward way in Wunderlich's approach, because it regards the content of semantic roles and not their relative prominence in the lambda-structure.

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 273 of 275 May 16, 2008 18:58

Another fact that is left unexplained is the subject behaviour of oblique experiencers, which was mentioned with special reference to Japanese at the end of section 17.3. The theta-structure of a two-place psychological predicate such as *understand* or *see* is λyλx P(x,y) in many approaches, including Wunderlich's. The corresponding prominence hierarchy is x > y with the experiencer as the highest role [−hr, +lr] and the target of sentience or cognition as the lowest role [+hr, −lr]. Wunderlich's structural linking principles explain the nominative or the initial basic position of the experiencer and the accusative or non-initial basic position of the target, a pattern that is found in English and other languages for such verbs. In order to capture an oblique experiencer, Wunderlich changes the prominence feature of the experiencer from [−hr] to [+hr]. This step licenses a lexical case in his approach. As an undesirable result, the structural information [−hr] is unavailable in syntax, although it is needed in order to explain at least those subject properties of oblique experiencers that are assumed to be structural in generative grammar. These properties include an initial basic position and the antecedent function for subject-bound anaphors.

In more general terms, lexical cases turn out to be more systematic than assumed by Wunderlich. Let us take a closer look at the semantic conditions for the dative in German (cf. Primus 1999: 65–74). They are formulated in a decompositional framework in terms of meaning components that are entailed or presupposed by the verb meaning. The first part of the condition is structural: the dative can only be assigned to the highest role, i.e. the proto-agent. The second qualification is non-structural: the participant in question is not the volitional causer of the situation denoted by the verb.[11] In accordance with these restrictions, the dative in German is licensed for the first argument *x* of the predicate types listed in (22):

(22) i) POSS(x,y), which is the relevant meaning component of verbs such as *give* (cf. (20a) above) and *gehören* 'own', cf. *mir gehört das Haus* 'I (DAT) own the house (NOM)'.

ii) SENTIENCE(x,y), which subsumes emotion and cognition as the relevant meaning component of verbs such as *sagen* 'tell', *zeigen* 'show' (cf. (20b) above), and *gefallen* 'like' cf. *mir gefällt der Film* 'I (DAT) like the film (NOM)'.

iii) Interactive predicates, which presuppose an activity or state in which the dative argument is involved as the highest role. This presupposed meaning component is mentioned after a semicolon in the translations, for convenience. Such verbs are *helfen* 'help *x*; *x* is doing or planning to do something', *danken* 'thank *x*; *x* has done something', *nachlaufen* 'run after *x*; *x* is also running', *zuwinken* 'wave to *x*; *x* perceives the waver'.

[11] This condition is necessary, but not sufficient: not all roles of this type occur in the dative, but except for a few verbs such as *entspringen* 'come from, originate', all dative verbal arguments satisfy this condition.

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 274 of 275 May 16, 2008 18:58

iv) Verbs whose nominative argument denotes an event the dative participant is involved in, e.g. *mir gelang der Sprung* 'I (DAT) managed to jump' and *mir passierte ein Unfall* 'I (DAT) had an accident'.
v) Symmetrical predicates such as *ähneln* and *gleichen* 'resemble, be similar'. Note that SIMILAR(y,x) entails SIMILAR(x,y), so that *x* is the highest role in the entailed predicate.

All these datives are licensed by a common theta-structural property, but this property cannot be expressed in the lambda-structure of all the predicate-types listed in (22). A pertinent common behaviour of datives in German is that under further conditions, they turn into subjects in the dative-passive construction: *ich* (NOM) *bekam Rosen* (ACC) *geschenkt* lit. 'I got roses given'. The dative-passive is not confined to the ditransitive verbs that license a structural case in Wunderlich's framework. It is also used with some of the verbs that select a lexical dative according to Wunderlich, e.g. *ich bekam geholfen* / *widersprochen* lit. 'I got helped / contradicted'.

17.5 Summary

All approaches that are concerned with the relationship between semantic roles and grammatical relations are able to capture the argument realization of transitive verbs selecting highly potent agents and strongly affected patients such as *break*, *open*, or *hit* in accusative languages. Approaches using role lists instead of semantic decompositions lack the means to cope with the large number of individual roles that are selected by the full range of verbs and with the reverse case pattern in ergative constructions. They also fail to offer a substantive explanation for basic questions pertaining to argument realization, such as: When do two or more roles count as identical for linking? Why are roles arrayed semantically or syntactically in a particular way?

Decompositional approaches fare much better. They explain role asymmetries on semantic grounds and reduce the inventory of general role concepts drastically without neglecting finer distinctions. Within this tradition, Wunderlich's structural approach links structural cases or syntactic relations to prominence relations in the semantic structure of predicates. Arguments bearing the same prominence relation, such as recipients and addressees of ditransitive verbs, are predicted to be linked to the same structural case or syntactic relation. However, purely structural semantic approaches such as Wunderlich's fail to incorporate finer distinctions of meaning that characterize the content of semantic roles and that are needed in order to explain a broader range of data.

978–0–19–920647–6 18-Spencer-ch17 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 275 of 275 May 16, 2008 18:58

The other kind of decompositional approach formulates generalizations in terms of semantic primitives such as volitionality, sentience, movement, and affectedness. Dowty's approach was discussed as a representative of this line of research. In such approaches, the linking principles are sensitive to the higher or lower number of semantic basic properties accumulated by an argument, e.g. volitionality, movement, and sentience vs. sentience alone. Such principles are able to explain the basic facts pertaining to the ergative–accusative typology and to morphological split intransitivity as well as the case variation of semantically underspecified roles in a straightforward way. But generalizations that hold irrespective of the number and type of proto-role properties remain unexplained.

In sum, the different decompositional views supplement each other as they focus on different dimensions of role semantics. There is preliminary evidence suggesting that these different dimensions correlate with different means of expressing grammatical relations: structural syntactic relations strongly correlate with structural role-semantic information, while case marking is particularly sensitive to the type and degree of involvement of a participant.

CHAPTER 18

SYNTACTIC EFFECTS OF MORPHOLOGICAL CASE

AD NEELEMAN
FRED WEERMAN

18.1 Introduction

It is a long-standing though controversial claim that morphological case has syntactic effects (see, for instance, Meillet 1949, 1950). In particular, word order freedom has been argued to be dependent on the presence of overt case markers. Thus, Latin and Classical Greek have both a rich case system and very free word order, while languages like Dutch and English lack both. The claim is controversial because there are languages that seem to have free word order but whose case system seems to be as deflected as that of Dutch and English. At an observational level the loss of morphological case from Latin to Italian, for example, has not led to a loss of free word order.

Counterexamples of this type seem to reduce the potential universal to a mere statistical correlation. Such correlations are of course not devoid of interest, but they require an explanation of a different type from absolute restrictions. The fact that the object *must* follow the verb in English, for instance, is a matter of

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 277 of 289 May 16, 2008 18:58

grammar. In contrast, if morphological case and free word order are related only statistically, one would probably appeal to more functional factors to account for the correlation. One could speculate, for example, that morphological case makes it easier to parse a language with free word order and that, as a consequence, free word order is found more often in languages with morphological case.

If, on the other hand, the effects of morphological case are encoded in the grammar, one of the implications is that languages with and without morphological case differ not only in their morphology but also in their syntax, contra standard versions of GB theory and minimalism.

The decision of whether a generalization requires a grammatical or a functional account cannot be taken on the basis of mere observation. The word order freedom found in Italian, for example, is partially dependent on clitic left dislocation (see Rizzi 1997 and the references mentioned there), while in Latin this is not true. In other words, it is possible that in the history of Italian word order freedom as a result of morphological case was lost, while word order freedom as a result of clitic left dislocation was acquired (or maintained).

This example emphasizes the importance of syntactic analysis, but it is equally important to analyse the morphological properties of the languages under discussion. After all, there is not a simple distinction between languages with and without morphological case. Rather, there is a wide variety of case systems, some richer and some poorer, and it may well be that certain syntactic effects only show up if the paradigm makes sufficient distinctions. In addition, it would be naive to equate morphological case in the intended sense with overt case. This is because the logic of a morphological analysis may imply the existence of case morphemes whose spell-out is zero. The absence in Icelandic of an overt accusative ending in the first masculine declension singular, for instance, does not entail that no accusative case exists in this part of the paradigm.

Conversely, the mere existence of a formal opposition in arguments does not imply the language in question has morphological case. A well-known example is the distinction between subject and object pronouns in languages whose regular nouns do not express nominative and accusative case. In general pronouns in these languages do not behave as if they have morphological case (see Emonds 1985 and Hudson 1995). Indeed, Weerman and Evers-Vermeul (2002) provide an alternative analysis of the opposition between subject and object pronouns that does not rely on morphological case. In view of this, we conjecture that, in order for syntactic effects to obtain, a language must possess a general paradigm that expresses case distinctions (in the sense of Pinker 1984); word-specific paradigms are not enough.

In this chapter, we explore a number of potential effects of morphological case that have to be encoded in the grammar. Space limitations prevent us from developing analyses of these effects; all we can do is discuss generalizations that we think are good candidates for phenomena that require a grammatical explanation. Alongside word order effects (section 18.2) we will also consider generalizations that

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 278 of 289 May 16, 2008 18:58

have to do with the form of constituents with particular grammatical functions (section 18.3). For example, quirky subjects (overt subjects that fail to agree with the finite verb in a structure that otherwise has subject–verb agreement) are found exclusively in languages with morphological case. Some general conclusions are given in section 18.4.

18.2 Effects on word order

We begin our discussion by considering a language without morphological case. Modern Dutch allows free ordering of arguments and adjunct. In (1) for example an adverb like *gisteren* 'yesterday' can be inserted in any of the marked positions (we use non-root environments to illustrate our claims about Dutch in order to avoid complications with verb second). However, the order of DP arguments is fixed (see 2a,b), as is the position of DP arguments with respect to the verb (see 2c). Moreover, separation of parts of arguments is rather limited (see 2d).

(1) *dat* Δ *Jan* Δ *Marie* Δ *de oude boeken* Δ *gaf*
that John Mary the old books gave
'that John gave the old books to Mary'

(2) a. **dat Jan de oude boeken Marie gaf*
that John the old books Mary gave
b. **dat de oude boeken Jan Marie gaf*
that the old books John Mary gave
c. **dat Jan Marie gaf de oude boeken*
that John Mary gave the old books
d. **dat Jan oude Marie boeken gaf*
that John old Mary books gave

With the exception of (2d) these restrictions do not hold of sentences in which the displaced constituent is interpreted and phonologically marked as a contrastive topic or contrastive focus (see Neeleman 1994 and Neeleman and van de Koot 2007; Haider and Rosengren 1998 and Frey 2001 discuss comparable phenomena in German). It can be argued that these interpretations license A′-movement. In (3), for example, movement of 'such books' is allowed if this constituent is interpreted as contrastive topic and marked by the so-called B-accent (Jackendoff 1972).

(3) *dat* [*zulke boeken*]$_i$ *zelfs Jan Marie niet* t_i *geeft*
that such books even John Mary not gives
'that such books even John does not give to Mary'

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 279 of 289 May 16, 2008 18:58

The conclusion that scrambling across adjuncts syntactically differs from scrambling across arguments is apparent from the fact that the former can create new binders for anaphors (see Vanden Wyngaerd 1989), while the latter, at least in Dutch, does not have this property (see Neeleman 1994). There is a sharp contrast between (4a–b) (in (4b) the fronted object 'the winners' is supposed to be a contrastive topic carrying the B-accent).

(4) a. *dat Jan de winnaars namens elkaar feliciteert*
that John the winners on-behalf-of each-other congratulates
'that John congratulates each winner on behalf of the other'
b. **dat de winnaars zelfs elkaars coach geen hand geeft*
that the winners even each-other's coach no hand gives
'that even the coach of one winner would not shake hands with the other winners'

As we will now show there is evidence that languages with a morphological case system that is rich enough allow structures like those in (2), even if no A′-movement is involved. We begin by considering scrambling across arguments.

German, as opposed to Dutch, has a dative–accusative distinction. The sentences in (5) show that it allows scrambling of accusative objects across dative objects (see Lenerz 1977). This does not require an interpretation as contrastive topic or focus and therefore modifiers like 'even' and 'only', which contribute to the grammaticality of (3), need not be inserted in (5b).

(5) a. *dass die Frau einem Mädchen einen Jungen vorstellt*
that the-NOM woman a-DAT girl-DAT a-ACC boy-ACC introduces
b. *dass die Frau einen Jungen einem Mädchen vorstellt*
that the-NOM woman a-ACC boy-ACC a-DAT girl-DAT introduces
'that the woman introduces a boy to a girl'

It is not completely clear whether the kind of word order freedom found in OV languages like German is also present in VO languages with morphological case. In Icelandic, for example, dative and accusative objects surface in both orders, but whether the operation responsible for this has the same syntactic signature as scrambling in Dutch and German is a matter of debate (see Collins and Thráinsson 1996).

What *is* clear, however, is that scrambling across arguments is permitted in other OV languages with morphological case. Middle Dutch and Old English are like German in marking accusative and dative case and allowing reordering of direct and indirect objects (see Weerman 1989). The context in which the relevant examples are found gives no indication that the displaced object must be a contrastive topic or focus. Some examples from thirteenth-century Middle Dutch prose texts are given in (6) (see Van Gestel et al. 1992).

(6) a. *so ghelove ic [...] den Grave die lettren [...] weder te ghevene*
thus promise I the count the letters back to give
'thus I promise to give back the letters to the count'
b. *[...] ic mine lettren [...] den vorseiden hospitale hebbe*
I my letters the above-mentioned hospital have
ghegheuven
given
'I gave my letters to the above-mentioned hospital'

For Old English, the word order in double object constructions has been described in much more detail than for Middle Dutch (see Koopman 1990, 1993). As expected, objects can appear in either order.

(7) a. *þæt ge Ongolþeode ætgædre mid us Drihtnes word*
that you the.English.DAT together with us God's WORD.ACC
bodige
preach
'that you will preach God's word together with us to the English'
b. *Ac gif we þa mirran gode gastlice geoffriað*
but if we then myrrh.ACC God.DAT spiritually offer
'but if we offer myrrh to God spiritually'

If (5b) is not derived by A′-movement, we expect it to be possible for the accusative argument to bind into the dative argument. This is indeed the case (see Bayer and Kornfilt 1994 and references mentioned there). The ungrammaticality of (8a) can be remedied by placing the direct object on the left of the indirect object.

(8) a. **weil ich [einigen Freunden von einander* $_1$*] [die Gäste]*$_1$
because I some.DAT friends of each other the.ACC guests
vorgestellt habe
introduced have
b. *weil ich [die Gäste]*$_1$ *[einigen Freunde von einander* $_1$*]*
because I the.ACC guests some.DAT friends of each other
vorgestellt habe
introduced have
'since I have introduced each guest to some friends of the others'

A similar observation can be made for Japanese. Objects that carry the accusative marker *-o* can be scrambled across subjects marked by *-ga*. Again, as opposed to what we have seen for Dutch, the displaced constituent does not require an interpretation as contrastive topic or focus (Junji Hamamatsu, personal communication).

(9) a. *John-ga subete-no hon-o yon-da*
John-NOM every-GEN book-ACC read-PAST

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 281 of 289 May 16, 2008 18:58

b. *Subete-no hon-o John-ga yon-da*
every-GEN book-ACC John-NOM read-PAST
'John read every book'

We expect OSV order to license binding by the object into the subject, a possibility illustrated in (10) (see Saito 1992 and Grewendorf and Sabel 1999, amongst others).

(10) a. ?* *Otagai-no sensei-ga gakusei-o sikatta*
each other-GEN teacher-NOM them-ACC scolded
b. *Gakusei-o otagai-no sensei-ga sikatta.*
students-ACC each-other-GEN teachers-NOM scolded
Lit: 'Each teacher scolded the other teacher's students.'

That morphological case is indeed responsible for scrambling across arguments is confirmed by the fact that the ordering of arguments that carry the same case must be preserved, even in languages that otherwise allow word order variation (a phenomenon sometimes referred to as 'word order freezing'). Certain verbs in German select two accusative objects and these verbs have been reported not to permit inversion of their objects (abstracting away from A′-scrambling; see Haider and Rosengren 1998 for discussion):

(11) a. *dass der Lehrer die Schüler diese Sprache lehrt*
that the teacher the pupils-ACC this-ACC language teaches
b. **dass der Lehrer diese Sprache die Schüler lehrt*
that the teacher this-ACC language the ACC-pupils teaches

As Jonathan Bobaljik (p.c.) points out, there are only few double accusative verbs in German, reducing the weight of the argument. However, a similar point can be made about Icelandic double dative verbs, which are more frequent. Moreover, Japanese has constructions in which both object and subject are marked by *-ga* and in these constructions the subject consistently precedes the object (see Vermeulen 2005):

(12) a. *John-ga nihongo-ga wakar-u.*
John-NOM Japanese-NOM understand-PRES
'John understands Japanese.'
b. **Nihongo-ga John-ga wakar-u.*
Japanese-NOM John-NOM understand-PRES

Another type of argument supporting the link between morphological case and word order freedom is diachronic in nature. If a process of deflection neutralizes the distinction between two cases, then scrambling between arguments carrying those cases should disappear. This seems to be correct for the Germanic OV languages

(although there are confounding factors in the history of English, associated with the change from OV to VO, which make it hard to test the prediction for this language).

There is an important distinction between German and Japanese: the patterns found in German also exist in Japanese, but the Japanese data have no parallel in German. That is, German OSV structures are always derived by A′-movement and consequently never allow binding by the object into the subject (see Grewendorf and Sabel 1999). This confirms the claim that languages need not have all the effects of morphological case or none. The central theoretical question that presents itself is what effects are conditioned by what aspects of the case system. A promising generalization for the data discussed so far is that scrambling across arguments is allowed as long as the argument appearing unexpectedly low in the structure carries a formal marker that identifies it (see Neeleman and Van de Koot 2006 for discussion of this and related data, such as the fixed word order with accusative–dative verbs). Since Japanese has a formal nominative marker, scrambling across the subject is possible. The German case system does not allow this, but the presence of dative in German allows scrambling across the indirect object. In a highly deflected language like Dutch, no scrambling across arguments is possible as no identifying morphological markers exist.

The above discussion suggests that scrambling across arguments can be licensed by morphological case, allowing constructions otherwise ruled out, as (2a–b) show. We now turn to a second type of word order freedom potentially linked to morphological case, namely extraposition of DP arguments in OV languages. As (2c) shows, Dutch does not allow this kind of extraposition, although it does allow extraposition of arguments that do not rely on the verb for case (PPs and both finite and non-finite CPs). As in the case of scrambling there is a confounding factor, namely rightward A′-movement of case-marked DPs (Heavy-NP Shift). In what follows, we therefore ignore processes of extraposition that have the signature of this operation.

The situation in Dutch is representative of other languages that lack morphological case: these typically do not allow extraposition of DPs, other than by Heavy-NP Shift (see Weerman 1989 and Neeleman and Weerman 1999). Earlier stages of Dutch, however, still had productive morphological case and, in our view as a result of this, allowed the relevant type of extraposition. Some examples are given below; the first two show extraposition of a single DP, the third shows that extraposition of multiple DPs is possible as well (as expected order between the extraposed DPs is free, see the references just mentioned):

(13) a. *Ic sal senden minen ingel vor dijn anschin*
I shall send my-ACC angel before your face
b. *hi soude dorbreken den muur*
he would through-break the-ACC wall

c. *sal dan v vader ... nit gheuen goede ghichten den ghenen*
will then your father not give good gifts-ACC those-DAT
die hem bidden?
that him ask

The option of DP extraposition disappears when the case system collapses, strengthening the correlation between the two. The Middle Dutch data have a parallel in Old English and Old High German (see Van Kemenade 1987 and Lenerz 1984, respectively). In general, extraposition seems to require morphological case. This generalization can be understood if morphological case allows arguments to escape licensing conditions that force case-less arguments to surface in position governed by the verb (in GB terminology). Notice that morphological case is not a sufficient condition: there are OV languages with morphological case that lack extraposition. Japanese and German are examples. To some extent, one can argue that this is the result of independent factors, such as rightward verb movement (for which there is evidence in Japanese). In other cases, it might be due to partial deflection (as we have argued for German in Neeleman and Weerman 1999).

The suggestion that extraposition is conditioned by morphological case provides an explanation for a well-known but ill-understood phenomenon. In languages with so-called differential object marking (such as Turkish, Sakha, Persian, Korean, and Japanese), objects may either show up with accusative case or without any overt marking. Abstracting away from various complications, the absence of overt case has the effect that the argument is interpreted as new, while the presence of accusative leads to an interpretation as discourse-anaphoric. This interpretive effect might provide a foothold for an explanation of the tendency of case-less object DPs to surface in a position immediately preceding the verb. In many languages, DPs that are discourse-anaphoric shift leftward, while DPs that represent new information surface in the rightmost position available (we are aware of the extensive literature of interpretive effects of scrambling, but space limitations prevent us from elaborating on the above characterization; see Neeleman and Van de Koot 2007 for some discussion). Of course, these generalizations are not without exceptions, but the observed pattern seems entirely consistent with them. Below, we illustrate how case omission bears on syntactic position for Sakha (all data from Van de Visser 2006: 77–81).

(14) a. *kini jabloko-*(nu) bügün sii-r*
s/he apple-(ACC) today eats
b. *kini bügün jabloko-(nu) sii-r*
s/he today apple-(ACC) eats

In languages that allow DP extraposition and that have differential object marking, the only DPs that can be extraposed are those that have an overt case marker, as shown in (15) for Sakha. This is unexpected from the perspective of information packaging (since cross-linguistically an interpretation as new tends to drive DPs

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 284 of 289 May 16, 2008 18:58

rightward), but follows directly from the idea that extraposition requires morphological case. (Note that the omission of a case ending differs from a zero morpheme whose existence is implied by the logic of a paradigm. Here, the logic of the paradigm excludes the existence of a zero morpheme.)

(15) *kini bügün sii-r jabloko-*(nu)*
s/he today eats apple-(ACC)

We now turn to a third potential effect of morphological case on word order, namely the separation of adjectives and nouns. The sentence in (2d) illustrates that this phenomenon does not exist in Dutch, and in fact, it has a limited distribution across languages. It seems that at least two conditions are relevant. The language must not have determiners (see Bošković 2005) and there must be agreement in case between the adjective and the noun (this is part of a general tendency for discontinuous configurations to favour agreement; see Bhat 1994). In languages such as Russian, Serbo-Croatian, Latin, and Czech, both conditions are met, and indeed separation is allowed, as illustrated by the Russian example in (16) (from Van Gelderen 2003).

(16) *Krasnuju ja včera videl sobaku*
red.FEM.SG.ACC I yesterday saw dog.FEM.SG.ACC
'Yesterday I saw the red dog'

Japanese and Korean do not have determiners, but since there is no case agreement between adjectives and nouns, separation is impossible (see Van Gelderen 2003 for related discussion). Icelandic and German have case agreement, but since they also have determiners, nouns and adjectives must remain together.

The obligatory absence of determiners can possibly be derived from the semantics of these elements and of pronominal modification: unification of the sets denoted by noun and adjective is impossible once the NP has been closed off by a determiner. However this may be, crucial here is the effect of case agreement: this type of morphological marking might serve as an alternative way of encoding an interpretive dependency that would require a structural relation otherwise.

The generalizations discussed in this section all link the presence of morphological case to the availability of certain derivations, namely scrambling across arguments, extraposition, and separation of nouns and prenominal adjectives. The generalizations in question do not directly link morphological case to word order freedom at an observational level, which can often be derived through alternative derivations. In the case of scrambling across arguments and extraposition, these alternative derivations involve A′-movement. An alternative way of achieving the separation of adjectives and nouns would be to link both to an incorporated argument in a polysynthetic language. Although superficially similar, these alternative derivations have rather different syntactic characteristics. This reinforces the point that effects of morphological case must be evaluated against analyses rather than mere data.

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 285 of 289 May 16, 2008 18:58

18.3 Effects on Form

The distribution of arguments is affected by morphological case, not only with respect to word order, but also with respect to the type of positions in which they are licensed. As is well-known the distribution of DP arguments in a language like English is severely limited. They cannot, for instance, show up as the complements of nouns or adjectives (see 17a,b). If they occur in the subject position of a finite sentence they must agree with the verb in structures where subject–verb agreement is available (see 17c) and they cannot be associated with a secondary predicate that they do not c-command (see (17d) and Williams 1980).

(17) a. He witnessed the destruction *(of) Carthage
b. He is proud *(of) his children
c. *The girls longs for world peace
d. *He put the meat [into [the oven]$_i$] hot$_i$

In all these cases the presence of morphological case may remedy the problem, as we will demonstrate now.

First, languages that have genitive case allow DPs as the complement of nouns without the help of a prepositional licenser. This well-known fact is illustrated in (18) for German and Japanese (the Japanese example is from Takahashi 1994: 398).

(18) a. *Das Haus meiner Freunde ist niedergebrannt*
the house my-GEN friends is down-burned
'My friends house has burned down'
b. *Yamaoka-sisyaku-no bessoo-ga go-rippa-da*
Yamaoka-viscount-GEN villa-NOM SH-splendour-COP
'It is Viscount Yamaoka whose villa is splendid'

Similarly, in the older stages of the Germanic languages the availability of genitive case allows nouns to take DP complements. In the process of deflection this option disappears and simultaneously a case-marking dummy preposition is introduced (*of* in English). Crucially there is no stage in which both genitive case and the dummy preposition are absent, suggesting a fairly strict grammatical relation. It seems to us that these observations may well be captured by the generalization that morphological case allows arguments to escape licensing conditions that force caseless arguments to surface in governed positions (see Neeleman and Weerman 1999).

Note that, as before, we are not dealing with a surface generalization; the existence of the construct state in Semitic shows that a concatenation of two DPs is possible even if no morphological case marking is present. Crucially, the construct state has very different properties from regular nominal complementation (see Borer 1989, among others).

Similar observations (and hence a similar explanation) hold for complements of adjectives (see Van Riemsdijk 1983). In the Germanic languages the effect seems

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 286 of 289 May 16, 2008 18:58

to be absolute for complements that follow the adjective: these must either have morphological case (as in 19), or be accompanied by a preposition (as in 17b). A limited number of adjectives are able to license complements that precede them, even if these do not bear morphological case. Although an analysis of these data is still pending, the general trend is the same as for complements of nouns.

(19) *Nochtan was hi onweger sijns lijfs*
nevertheless was he indifferent his-GEN body-GEN
'Nevertheless he did not mind his body'

That morphological case allows arguments to appear in positions where they are not otherwise licensed is confirmed by the distribution of so-called quirky subjects. We have already seen that in languages without morphological case subjects must agree with the finite verb where subject–verb agreement is available (see 17c). This is presumably because agreement is a precondition for the licensing of nominative arguments (which in the languages in question are plausibly analysed as bare DPs, as opposed to KPs; see Weerman 1989, Bittner and Hale 1996a, and Neeleman and Weerman 1999).

In at least some languages with morphological case, non-agreeing subjects can appear in configurations that otherwise require agreement, as long as they bear a case other than nominative. This is presumably because these other cases are not licensed by agreement. That the relevant non-agreeing DPs are indeed subjects is shown by a number of tests (see Zaenen, Maling, and Thráinsson 1985). Some Middle Dutch and Icelandic examples can be given in (20a) and (20 b–c), respectively. The finite verb in examples of this type always surfaces in the default third person singular form (unless there is agreement with a lower argument).

(20) a. *Doe gedachte den Sassen der scaden*
then thought the Saxons-DAT the damage-GEN
b. *Strakunum leiddist í skólann*
the.boys.DAT got.bored in the.school
c. *Strákana vantað í skólann*
the.boys.ACC lacked in the.school

As far as we know, all languages that allow non-agreeing subjects in configurations that otherwise require agreement have morphological case. Relevant examples of non-agreeing subjects can be found in Latin, Old French, and Japanese, and the older stages of the Germanic languages. (N.B. This generalization abstracts away from agreement-reducing rules that apply in the context of inversion, as attested in Modern Standard Arabic and various other languages.)

The reverse generalization does not hold. Modern German and Russian, for instance, do not allow quirky subjects, although they do have morphological case. As before, one can try to explain the asymmetric nature of the generalization in terms of independent factors. One of these might be a certain degree of deflection (as

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 287 of 289 May 16, 2008 18:58

found in German). It is certainly true that the disappearance of morphological case correlates with the disappearance of quirky subjects in closely associated languages such as Dutch. Another possibility is that the verbal agreement system of a language does not allow the use of default forms where more specific forms are in principle available.

The cross-linguistic distribution of quirky subjects nicely illustrates a point we have made in the introduction, namely that pronominal formal distinctions cannot be equated to morphological case: so-called accusative pronouns in languages like Dutch and English can never be used as quirky subjects.

The observation that quirky subjects must have morphological case might have a parallel in the tendency, noted by Trask (1979), for ergative case to be marked morphologically. Interestingly, if only one argument agrees in ergative languages, there is a very strong tendency for this to be the absolutive rather than the ergative DP. If ergative DPs do agree, the relevant agreement affix tends to be subordinated to the absolutive agreement affix (see Bittner and Hale 1996b). Thus, the link between agreement and absolutive case seems much weaker than that between agreement and ergative case. One might consequently attempt an analysis of ergative subjects as a special instance of quirky subjects. On this view, ergative case would be a systematized quirky case. We admit that this is a direction of research rather than a firm result, because the typological generalizations on which the suggestion is based are not absolute.

As already suggested in the previous section, morphologically expressed agreement in case between an adjective and a nominal constituent can allow an interpretation of the two as associated even if the usual structural configuration is not present. Thus case agreement licenses fronting of prenominal adjectives (see 16). We may expect the same kind of effect with secondary predicates. The usual condition of c-command, illustrated in (17d), need not hold when the secondary predicate shows case agreement with its subject. Notice that here we do not expect the presence of determiners to cause any problems, as secondary predicates are interpreted after the DP has been closed off.

Maling (2001) argues that case agreement can indeed make up for an absence of c-command between a secondary predicate and its subject (see also Hale 1981, Laughren 1992, Rapoport 1991: 179ff.). Maling cites Icelandic as a case in point:

(21) a. *Ég hrifsaiði af honum nestispakkann glorsoltnum og þeytti*
I snatched from him.DAT his.lunchbox ravenous.DAT and threw
honum í sjóinn
it into the.sea

b. *Settu kjötið inn í ofninn eldheitan*
put the.meat into the.oven.ACC redhot.ACC

That case agreement is indeed the conditional factor for secondary predication in the absence of c-command is confirmed by several observations. First, languages

that have phi-feature agreement but no case agreement do not allow constructions parallel to those in (21). Norwegian and Italian are examples. Second, in languages that have invariant case endings on depictives secondary predication requires c-command. This can be shown in Russian, where depictives are consistently marked with instrumental case.

The data discussed so far show that morphological case can license arguments in positions they may not otherwise occupy. Thus more is possible in the presence of morphological case. The reverse type of effect exists as well. To give just one example, again borrowed from Maling 2001, dative case blocks topic drop in German (see also Sternefeld 1985). In a closely associated language like Dutch, which does not have a dative–accusative opposition, topic drop is allowed in parallel examples:

(22) a. *Was ist denn mit dem Peter?*
what is then with the.DAT Peter
*(*Dem*) *hab' ich schon geholfen*
(him.DAT) have I already helped

b. *Hoe zit het eigenlijk met Peter?*
how sits it really with Peter
(*Die*) *heb ik al geholpen*
(him) have I already helped

18.4 Concluding remarks

If there is a connection between morphological case and syntax, this does not only make typological and diachronic predictions of the sort discussed here, but also predictions about language acquisition. In particular, we do not expect the child to produce structures dependent on morphological case if morphological case has not been acquired yet. At present, we cannot evaluate this prediction, although there is some evidence in the literature that it might be correct. Eisenbeiss (1994b) shows that scrambling across arguments does not appear in German prior to the introduction of the morphological case system. Similarly, Slobin (1966, 1982) claims that case is acquired relatively late in Russian and relatively early in Turkish. As a consequence of this, Russian children begin with a rigid word order, while free word order effects already exist in the early stages of Turkish.

Let us finally consider some general theoretical consequences of the existence of syntactic effects of morphological case. It is sometimes argued that models like distributed morphology (see Halle and Marantz 1993) are incompatible with such effects, because they take morphology to be ordered after the syntax. To our mind, this is an incorrect conclusion. What is crucial to explain syntactic effects

978–0–19–920647–6 19-Spencer-ch18 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 289 of 289 May 16, 2008 18:58

of morphological case is that in the process of acquisition the presence of case distinctions in the input leads to a richer set of syntactic case features, while the absence of such distinctions leads to fewer features. There is nothing inherent to distributed morphology that forbids variation in the set of syntactic features.

If our take on the data is correct, the phenomena described above do, of course, argue against any theory that assumes that the same set of features underlies the syntax of all languages (such as standard versions of GB theory and minimalism). The inventory of features must be subject to parameterization. Some authors have developed proposals in which the parameterization is rather drastic in that different licensing mechanisms are proposed for arguments that have morphologically expressed case and arguments that do not (see Bouchard 1996, 1998, Kiparsky 1997). We believe that this approach is overly pessimistic and that there is room for a model in which the effects of the same set of principles vary depending on what features exist in a particular language (see Neeleman and Weerman 1999).

CHAPTER 19

CASE AND ALTERNATIVE STRATEGIES

WORD ORDER AND AGREEMENT MARKING

ANNA SIEWIERSKA
DIK BAKKER

19.1 Introduction

It is often assumed and sometimes explicitly stated that both agreement marking and word order constitute viable alternatives to morphological case with respect to some subset of the functions that case marking may fulfil. This chapter will explore to what extent this is indeed so and how the three forms of marking interact with each other on a cross-linguistic basis.

The discussion will begin in section 19.2 with an overview of what are typically considered to be the primary functions of case marking, agreement, and word order, concentrating on the areas of overlap between case marking and the other two forms of morphosyntactic encoding. Then we will proceed to consider in more detail the ways in which case marking interacts with word order and agreement.

Section 19.3 will be devoted to a consideration of the relationship between case marking and basic clausal constituent order first noted by Greenberg (1963b), namely the predilection for case marking of core grammatical relations in languages with basic APV order and the scarcity of case marking in languages with basic AVP order. The existence of the above preferences will be explored from the perspective of the potential advantages for on-line processing stemming from the earliest correct recognition of structures, as elaborated most extensively in Hawkins (2004). In section 19.4 we will turn to a consideration of the degree of complementarity and overlap in the overt marking of case and agreement language-internally, concentrating again on core grammatical relations including those found in ditransitive clauses. The discussion will focus on the differences in the degree of overlap in case and agreement marking exhibited by the verbal arguments. In section 19.5 we will close the discussion with some remarks on the relationship between case and agreement marking and word order flexibility.

19.2 The functions of case marking, agreement, and word order

The primary function of case marking, be it via affix or adposition, is typically seen to be a relational one, namely of denoting the nature of the semantic dependency obtaining between the verb and its less predictable dependents, i.e. its adjuncts (see e.g. Moravcsik 1974; Lehmann 1988; Croft 1988). Case identifies the dependent leaving implicit the head and simultaneously indicates the nature of the semantic relation of the adjunct. The more unpredictable the semantic nature of the dependency relation, the more likely it is to be overtly marked by case. By the same token, arguments, in contrast to adjuncts, are much less likely to bear case since the nature of the dependency relation between a head and its dependent arguments is largely predictable from the lexical properties of the head and less often, from the properties of the arguments. In addition to this primary relational function associated with adjuncts, case marking may also perform two other functions, which concern essentially arguments only. The first of these is the discriminatory or differentiating function (Comrie 1978, 1989 and Dixon 1979, 1994), i.e. the use of case to distinguish the A from the P in transitive clauses and the R (recipient) from the T (theme) in ditransitive ones. This role of case marking rests essentially on the presence vs. absence of marking of an argument as opposed to the nature of the case marking. Since what is at stake is mere differentiation of one argument from another, only minimal case marking of core arguments is required. For example, in transitive clauses only the A or the P needs to be case-marked rather than

both. Furthermore, in terms of discrimination no preference is assigned to which argument should receive overt case marking. Thus the overt marking of just the P in accusative alignment, as in the Brazilian language Kwaza (1) or of just the A in ergative alignment, as in the Tibeto-Burman language Manange (2) are functionally equivalent.

(1) Kwaza (van der Voort 2000: 57)
wa zjwau-'wa e'cyi-ki
bee Joao-AN.P sting-DECL
'A wasp stung Joao.'

(2) Manange (Hildebrandt 2004: 68)
/1/mriŋ-tse naka /2/phuŋ /2/khol-tsi
woman-ERG chicken egg boil-PRF
'The woman boiled the egg.'

The second additional function that case marking may fulfil is that of indexing properties of the referents of arguments or of the clause itself (see e.g. Silverstein 1976; Hopper and Thompson 1980). The properties of the referents of arguments may concern their inherent characteristics (e.g. animacy) or contingent ones (e.g. definiteness, referentiality, focus). Unlike in its discriminatory function, in its indexing function case marking is not driven by economy, that is, in transitive clauses both the A and P may bear overt case marking under appropriate circumstances. Moreover, if only one of two arguments bears case, it is highly relevant which of the arguments does. One line of reasoning associates case marking with inherent saliency as reflected on the person or animacy hierarchies. Another sees case marking as favouring arguments (and clauses) which display some departures from prototypicality be it with respect to the properties of their referents (inherent or discourse properties) or their morphosyntactic realization. An example of indexical case marking is the so-called differential case marking of the P, i.e. the use of case marking only on Ps that are pronominal, human, or animate and/or definite and/or specific. For instance, in the Colombian language Desano (Miller 1999: 57) the relevant conditioning seems to be specificity, as suggested by the presence of the clitic case marker *-re* in (3a) as compared to (3b) and (3c).

(3) Desano (Miller 1999: 78, 55, 67)
a. *bari-re ai-ga-ke*
food-SPC take-move-IMP
'Take the food to another spot.'
b. *era bere di?ta ba?ba-ba*
they fruit only eat-3PL
'They eat mere fruit.'
c. *su?ri koe-go ii-ku-bo pera-ge*
clothes wash-F.SG do-assume-3F.SG port-LOC
'She (probably) is washing clothes at the river landing.'

978–0–19–920647–6 20-Spencer-ch19 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 293 of 303 May 16, 2008 18:58

(For further examples of differential case marking see Malchukov and de Swart in this volume.)

Whereas case marking is primarily a relational encoding strategy denoting the relation holding between two entities, agreement is an indexing strategy denoting the properties of one of the entities in the agreement relationship (Croft 1988: 173; 2003: 199). Using the terminology introduced by Corbett (1983), we will refer to the entity whose properties are indexed as the controller and the entity on which the index is placed, the target of the agreement relation. Thus, for example, in the Desano example in (3b) the controller is the A argument *era* 'they', the target is the verb *ba?ba* 'eat' and the agreement index on the target is *-ba* '3PL'. The notion of agreement is variously conceived of in the literature. Our use of the term here encompasses both so-called grammatical agreement, where the controller and target are both overt in the same clause (or phrase), as in (3b), and anaphoric agreement where there is no overt controller in the clause (or phrase) featuring the target, as in (3c).[1] Further, whereas some scholars restrict the notion of agreement to indexing via affixal marking, for us it also includes marking via clitics. In what follows we will confine our attention to agreement involving person, which more often than not simultaneously encodes number and/or gender but not necessarily so. Person agreement is thus considered to be primarily a means of keeping track of referents in the discourse via their index of features. Under our view person agreement is thus a form of pronominalization used in preference to free personal pronouns for highly salient discourse referents, which the speaker assumes to be easily accessible to the addressee (see e.g. Lehmann 1988; Ariel 2000; Siewierska 2004). Since cognitively salient referents tend to be encoded as arguments rather than as adjuncts, agreement, unlike case marking, is primarily associated with arguments. It thus enters into competition with case marking only in regard to the secondary functions of the latter, that is with respect to discriminating between two arguments and the indexing of their special properties. As a means of discriminating between arguments, agreement marking is comparable to case marking only if the arguments in question differ in terms of the person, number, or gender features indexed by agreement. By way of illustration, consider, for instance, the clauses in (4) from Gumawana, an Oceanic language spoken in the Milne Bay Province of Papua New Guinea.

(4) Gumawana (Olson 1992: 327, 326)
   a. *Koloto vaniva-yao i-duduwe-di*
      man woman-PL 3SG-call:TR-3PL
      'The man called the women.'
   b. *Koloto-ya-di vavina si-duduwe-Ø*
      man-REF-3PL woman 3PL-call.TR.3SG
      'The men called the woman.'

[1] The terms grammatical and anaphoric agreement were introduced by Bresnan and Mchombo (1986). An elaboration of their typology is discussed in Siewierska (2004: 12–127).

c. *Topiyo Kelebi i-tala-i-Ø*
Topiyo Kelebi 3SG-hit-TR-3SG
'Topiyo hit Kelebi.'

As the above examples suggest, in Gumawana, agreement with the S or A is marked by prefixes and with the P by suffixes. Thus since in (4a) the prefix is 3SG and the suffix 3PL, *koloto* 'man' is identified as the A and *vanivayao* 'women' as the P, irrespective of any discriminatory role that the word order may play. And analogously in (4b). In (4c), by contrast, the agreement marking fulfils no differentiating function since both of the arguments are third person singular and if it were not for the word order, the clause would be ambiguous, that is, either Topiyo or Kelebi could be the A or the P. In view of the above, agreement marking is a less versatile and reliable means of differentiating between arguments than case marking. As for the indexing of additional semantic or pragmatic properties of arguments, agreement is just as good a means of such indexation as case marking. Moreover, it tends to be used to index the same range of properties as case marking and for the same type of arguments, i.e. for the P as opposed to the A. Thus, for example, in Gapapaiwa (McGuckin 2002: 300, 307, 309) – another Oceanic language of the Milne Bay Province of Papua New Guinea – agreement marking of the P occurs when the P is a specific higher animate (e.g. human, spirit, or pet) but not a lower animate or inanimate. Compare (5a) with the suffixal P agreement marking with (5b,c) which feature no such marker.

(5) Gapapaiwa (McGuckin 2002: 310, 305)
a. *Wivine-si a-vi-yava-i-si*
woman-3PL 1SG-CAUS.PAST-count-TR-3PL
'I counted the women.'
b. *aririta i-peyari sasara a-vunuwa*
aririta.bird 3:NPRS-many very 1SG-capture
'I caught a lot of aririta birds.'
c. *wakima a-vi-yava-i*
rock 1SG-CAUS.PAST-count-TR
'I counted rocks.'

A similar situation is found in Acehnese, Hua, Kairiru, Mundari, and Noon, while in languages such as Bulgarian, Greek, Persian, Rumanian, Spanish, and Tinrin the presence of P agreement depends on specificity or definiteness.

Turning to word order, its basic function is taken to be the sequencing of information in ways which best reflect the communicative intentions of the speaker and simultaneously enable these intentions to be successfully and speedily processed by the addressee. Whether this complex goal can be best achieved by linearizing the words in an utterance in terms of their informational load at a given point in the discourse as captured, for example, by Givón's (1988) Task Urgency Principle or in terms of human parsing preferences, as reflected in Hawkins' (1994, 2004) parsing

978–0–19–920647–6 20-Spencer-ch19 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 295 of 303 May 16, 2008 18:58

principles is not yet clear. While with respect to its primary function word order differs quite radically from case marking (especially under the parsing approach), there is some overlap between the two again in regard to argument differentiation and indexation. Consistent placement of arguments on opposite sides of the verb as in AVP or PVA order in particular, and to a lesser extent in a fixed sequence before the verb (e.g. APV, PAV) or after the verb (e.g. VAP, VPA) is a means of distinguishing arguments from each other. Recall, for example, the disambiguating function of the placement of the A before the P in the Gumawana example in (3c) in the absence of any contribution from the agreement markers on the verb. The placement of constituents in designated clausal locations may in turn be used as a means of indexing properties of their referents, in particular specificity (or lack thereof), referentiality, or focus. Thus, for example, immediate preverbal location is a common focus position in APV languages while immediate positioning after the verb in AVP ones may be suggestive of nonspecificity or incorporation. Significantly, unlike case marking or agreement marking, word order does not tend to be used to index the inherent properties of referents, i.e. animacy, humanness, or person. Nor does it tend to single out the P in preference to the A or S. In fact it is arguably the intransitive S that is most likely to exhibit variation in its clausal location.

We have seen that despite their different primary functions case marking, agreement, and word order may all be used to discriminate between arguments and index their properties. We have also seen that with respect to argument discrimination word order is a better alternative strategy than agreement, at least when it is relatively stable. With respect to referent indexing agreement is more comparable to case marking than word order both in terms of the nature of the features indexed and the arguments involved. This suggests that we may expect there to be a significant interaction between case marking and word order with respect to argument differentiation and also between case marking and agreement marking with respect to argument indexation. Let us now consider whether this is indeed so.

19.3 Case marking and basic order of the verbal arguments

Of the various relationships posited in the literature between case marking and basic clausal order (cf. Siewierska 1996), the best known is Greenberg's (1963b: 96) universal 41 which is: 'If in a language the verb follows both the nominal subject and nominal object as the dominant order, the language almost always has a case system.' Although there are in fact many APV languages which lack case marking

Table 19.1. Distribution of case over word order types

	V-final	V-medial	V-initial
Dryer N=502	72% (181/253)	14% (26/190)	47% (28/59)
S&B N=417	71% (143/202)	21% (32/150)	42% (27/65)

such as Gumawana exemplified in (4) cited earlier, cross-linguistic studies strongly confirm that APV languages are much more likely to exhibit case marking than languages manifesting other types of basic order.

This is quite evident in the data presented in Table 19.1, reflecting the distribution of case marking relative to basic clausal order expressed in terms of the three-way typology of verb-final, verb-medial, and verb-initial, with the APV and PAV languages grouped under verb-final, etc. The data relate to the 582 languages classified as manifesting a basic order in Dryer (2002) for which case information was available to the author and the 417 languages analysed as evidencing a basic order in our current sample. In both samples the variation in case marking relative to basic word order is highly statistically significant (at the 0.5% level). As the figures plainly show, this is due on the one hand to the predilection for case marking in verb-final languages, and on the other, to the dispreference for case marking in verb-medial languages.

The traditional explanation for the relationship between case marking and basic word order depicted in Table 19.1 runs as follows. The dispreference for case marking in verb-medial languages is attributed to the overlap in function of the two forms of morphosyntactic encoding. The placement of the A and P on opposite sides of the verb is seen to be just as good a means of discriminating between the transitive arguments as the overt case marking of either or both of them. Accordingly, once the A is separated from the P by the verb, the marking of either the A or the P via case, for purposes of discrimination, is considered to be superfluous. This line of argument is extended to account for the higher incidence of case marking in both verb-initial and verb-final languages as compared to verb-second ones. In both verb-initial and verb-final languages the placement of the A and P on the same side of the verb allows for the A and P to be differentiated from each other in terms of their linear order alone. However, linear order is a less robust form of argument discrimination than placement on opposite sides of the verb since it relies on the overt presence of the A and P and the absence of other clausal constituents which could interfere in distinguishing the A and the P from each other. Case marking thus ensures that the A and P are clearly identified. As for the considerably higher incidence of case marking in verb-final as opposed to verb-initial languages, this is to a large extent attributable to verb position. Given that the verb's arguments are largely predictable from the semantics of the verb, the placement of the verb in

initial position provides important clues about the nature of the A and P and thus facilitates the identification of the two. This reduces the need for case marking. In verb-final languages, by contrast, since the A and P occur before the verb, relying on the verb to differentiate the A and P from each other would delay the task of identification. Case marking avoids such a delay by allowing the A and P to be distinguished from each other before the verb is reached.

A somewhat more sophisticated version of the above account of the relationship between case marking and basic clausal order has been recently elaborated by Hawkins (2004) within the context of his processing model of grammatical structure. Taking as his point of departure the assumption that grammars have conventionalized structures in proportion to their degree of preference in performance, Hawkins posits several efficiency principles, which he sees as facilitating on-line processing. The one most relevant to the current discussion is Maximize On-line Processing (MaOP) which defines a preference for the earliest possible correct recognition of assignments of properties to forms. The MaOP predicts that correct recognition can be achieved earliest if morphosyntactic marking facilitating recognition and precluding misassignments and unassignments is skewed to the left of a clause. The positioning of morphosyntactic marking on the left-side of the clause translates into the presence of case marking on arguments when they occur before the verb, and the presence of agreement marking on the verb when they occur after the verb. Although both forms of morphosyntactic encoding, case marking, and agreement marking, are considered to be compatible with all word order types, verb-final languages are predicted as favouring case marking over agreement marking, verb-initial languages as favouring agreement marking over case marking.[2] Verb-medial languages in turn are seen to disfavour either form of morphosyntactic encoding by virtue of the disambiguating function of the verb position. Nonetheless, Hawkins (2004: 245) argues that the overall level of case marking should be lower than that of agreement marking since agreement of the verb with an argument which follows it is in line with the MaOP, while agreement with the preverbal A in so-called pro-drop languages satisfies another efficiency principle, namely form minimization.

The relationship between word order type and the other two forms of morphosyntactic encoding outlined above is actually formulated by Hawkins not in terms of case marking and agreement marking per se but rather with reference to what he calls rich case marking and rich agreement marking. The notion of rich case marking does not differ from what is typically understood by the term case marking as applied to lexical NPs in transitive clauses, that is, it covers the presence of affixal or adpositional marking on the A or P or both which enables the two to be distinguished from each other. Rich agreement marking is agreement by affix or

[2] That agreement in verb-initial languages carries processing advantages is widely accepted. See, for example, Nichols (1992: 108) or Siewierska and Bakker (1996).

Table 19.2. Distribution of rich case and agreement over word order types

	Rich Case		Rich Agreement	
	Dryer N=502	S&B N=417	Dryer N=582	S&B N=417
V-final	72% (181/253)	71% (143/202)	49% (140/283)	51% (104/202)
V-medial	14% (26/190)	21% (32/150)	44% (94/213)	54% (81/150)
V-initial	47% (28/59)	42% (27/65)	56% (48/86)	69% (45/65)

clitic with both the A and P rather than with just one or the other.[3] Thus verb-final languages are predicted to favour rich case marking over rich agreement marking, and verb-initial languages as favouring rich agreement over rich case with verb-medial languages exhibiting a weaker preference for rich agreement over rich case. In support of the predictions of the MaOP Hawkins cites data from the previously mentioned sample in Dryer (2002). These data together with ours are presented in Table 19.2.[4]

We see that in both samples the relevant predictions made by the MaOP are borne out. In verb-final languages rich case marking is more common than rich agreement marking (72% vs. 49% in Dryer's sample and 71% vs. 51% in ours) and in verb-initial languages rich agreement marking is more common than rich case marking (56% vs. 47% in Dryer's sample and 69% vs. 42% in ours). Further, in verb-medial languages rich agreement clearly prevails over rich case marking (44% vs. 14% in Dryer's sample and 54% vs. 21% in ours). It is important to note though that the distribution of rich agreement relative to word order type is much more uniform than that of rich case marking. Particularly telling is the close similarity in the levels of rich agreement marking in verb-medial and verb-final languages in both samples. Since only agreement with arguments that follow the verb as opposed to those that precede it is in line with the MaOP, one would expect verb-medial languages to evince a higher level of rich agreement than verb-final ones. Yet as the

[3] Strictly speaking the agreement in question should be grammatical rather than anaphoric in the sense of Bresnan and Mchombo (1986) since only the former co-occurs with lexical NPs. Hawkins does not, however, discuss the issue and the data he cites clearly embrace both types of agreement markers. Therefore our comparative data in Table 19.2 also refer to both types of agreement. In section 19.4, however, we will restrict ourselves to grammatical agreement only.

[4] It is important to note that Hawkins frames his predictions in regard to the distribution of case and agreement in relative not absolute terms. Furthermore, it should be noted that, although Nichols (1992) classifies several languages as having both head and dependent marking of all three arguments, in her classification not all of these markers need be overt. When we look at the number of head and dependent marking points assigned to the languages in her Appendix Two, it turns out that if a language has agreement affixes for the A, R, and T, then at least one of the corresponding lexical NPs bears no overt marker of case, as in Basque for example. And if all the lexical NPs bear overt case marking, then there are restrictions on the number of non-zero agreement markers on the verb, as in Georgian.

data in Table 19.2 reveal, this is only marginally so in our sample while in Dryer's quite the opposite obtains.

The distribution of case and agreement marking relative to word order type depicted in Table 19.2 tells us little about how the two forms of morphosyntactic encoding interact with each other *within* languages as opposed to *across* languages. Hawkins' dichotomizing approach with respect to the parsing advantages of case or agreement marking depending on word order type suggests that there should be a considerable complementarity in the distribution of these two types of marking within languages. On the other hand, if the parsing advantages of agreement are not as great as those of case marking (recall that agreement can disambiguate the A from the P only when the two differ in person, number, or gender) one may expect some overlap between the two rather than mutual exclusion particularly in verb-initial and verb-medial languages. Let us therefore take a closer look at the patterns of distribution of case and agreement marking within languages.

19.4 The language-internal distribution of case marking and agreement marking

As is well known, case marking of arguments is overall considerably less common cross-linguistically than agreement marking. Moreover, case marking is more often found in languages that also display agreement marking than in languages in which it is the only means of morphological argument encoding. In our sample, of the four logically possible combinations of the occurrence of case and agreement marking of the arguments in transitive clauses, the most common cross-linguistically is agreement alone (44%) followed by case and agreement (37%), then case alone (10%), and finally neither (8%). Thus, there is a considerable amount of complementarity in the distribution of case and agreement marking. The degree of overlap is substantial enough to warrant closer scrutiny. Let us therefore take a look at which arguments are actually overtly marked by case and which by agreement.

As discussed in section 19.2, in transitive clauses case may be overtly marked on the A or on the P or on both. The differentiation of the A from the P can be achieved by overt case marking of either of the two only. Indeed, the overt marking of both is less common than the overt marking of just the A or of just the P. Of the 213 languages with case marking in transitive clauses in our sample, only 41 (19%) exhibit overt case marking of both the A and P. In contrast to case marking, agreement with both the A and the P is cross-linguistically more common

than agreement with either the A or the P. Of the 316 languages in the sample which display grammatical agreement marking 182 (58%) manifest agreement with both the A and P. Furthermore, when only one argument displays agreement it is overwhelmingly the A (131; 41%) as opposed to the P (6; 2%) which does so. In the light of the above, any overlap in overt marking via case and agreement is more likely to involve the A than the P. And indeed, in our sample the degree of overlap in overt case and grammatical agreement marking of the A is 66% (81 out of 123 languages with case marking for the A) while of the P it is only 32% (42 out of 131).

In ditransitive clauses case marking favours the R over the T, with adpositional marking of the R being somewhat more common than affixal marking. The overt case marking of both the R and T occurs more frequently than of the two arguments of transitive clauses.[5] Agreement marking also favours the R over the T. Further, unlike in transitive clauses, agreement with both the R and T is strongly disfavoured over agreement with either the R or the T (Gensler 2002; Siewierska 2003; Haspelmath 2005b). Given that the R more often than the T exhibits agreement, it is the R which is the more likely of the two to exhibit overlap in overt case and agreement marking. Among the languages in our sample, overt marking of the R via case and agreement obtains in 7% (30/419) of the languages. Thus, double marking of the R is even less common than double marking of the P. In sum, the likelihood of an argument exhibiting both overt agreement and case marking conforms to the hierarchy in (7), with overlap in marking declining as we proceed from left to right.

(6) A > P > R

Since agreement is a property of arguments, what the hierarchy in (7) essentially reflects is the increasing sensitivity or dependence of agreement on the presence of overt case marking. The further down the argument prominence hierarchy we go, the more likely agreement is to be optional rather than obligatory. By the same token, the greater the sensitivity of agreement marking to the presence of overt case marking the less the likelihood that the two forms of marking will overlap. Agreement marking with the A is typically obligatory. It is very rarely conditioned by the presence of case marking alone. Virtually all instances of agreement with the A tied to case marking that we are aware of turn out on closer inspection to be either lexically determined (i.e. occur with verbs requiring experiencers rather than agents) or to involve tense or aspectual distinctions. For example, in Hindi/Urdu and various other related languages (see Klaiman 1987) the A evinces agreement in the imperfective when it bears no overt case marking (7a), but in the perfective where it bears the ergative case (7b), the agreement is with the P. (*RooTtii* 'bread' has feminine gender.)

[5] The relevant figures for languages in which both the A and P markers are grammatical as opposed to anaphoric in our sample are Vf = 41%, Vm = 33%, and Vi = 46%.

978–0–19–920647–6 20-Spencer-ch19 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 301 of 303 May 16, 2008 18:58

(7) Hindi (Comrie 1984: 858)

a. *laRiyaa rooTtii khaa-tii hai*
girls bread eat-IMPF.F.PL be.PRES.3PL
'The girls eat bread.'

b. *laRiyaa=ne rooTtii khaa-ii*
girls=ERG bread eat-PRF.F.SG
'The girls ate bread.'

c. *laRiyaa=ne rooTtii=koo khaa-yaa*
girls=ERG bread=ACC eat-PRF.M.SG
'The girls ate bread.'

In contrast to agreement with the A, agreement with the P is frequently not obligatory and more often at least partially conditioned by case marking. Moreover, case-conditioned variation in agreement with the P is generally associated not with tense or aspect but with the referential properties of the P. This can be observed also in Hindi/Urdu. As a comparison of (7b) with (7c) shows, Hindi/Urdu displays P agreement in the perfective but only when the P is not marked by the case clitic *koo.* When *koo* occurs, as in (7c), there is no agreement with either the P or the A. The verb displays only default masculine singular marking. The marker *koo*, however, is not just a case marker but also a marker of specificity. Recall from section 19.2 that agreement marking like case marking may fulfil an indexing function, with indexing features such as definiteness, specificity, humanness, or animacy. It may therefore be argued that the failure of agreement marking to occur in the presence of *koo* is due to matters of economy. More specifically, if neither case nor agreement is obligatory and thus either may potentially be used to index the properties of referents, it is uneconomical to signal the same property of a referent twice, once by case and once by agreement. That agreement should give way to case, rather than vice versa, may in turn be attributed to the fact that referential features such as specificity, as compared to inherent ones, are essentially properties of lexical arguments and as such are better marked on the arguments themselves rather than on the verb. Turning to the R, agreement with the R is considerably more frequently conditioned by case marking than agreement with either the A or the P. The case marking in question, however, tends to be adpositional rather than affixal. The presence of agreement with the R is not tied to differences in tense/aspect or referential distinctions. Rather what is generally at issue here is the argument status of the R. Adpositionally marked Rs are lower on the argument prominence hierarchy than those lacking case marking. And it is only the latter which tend to display agreement marking, as illustrated in (8) from Southern Tiwa. Note that (8b) manifests agreement marking with both the T and R.

(8) Southern Tiwa (Rosen 1990: 674)

a. *bi-musa-wia-ban ùide áy*
1SG.3PL-cat-give-PAST child to
'I gave the cats to the child.'

b. *Ùide tam-musa-wia-ban*
child 1SG.3PL.3SG-cat-give-PAST
'I gave the cats to the child.'

Adpositionally marked Rs only exceptionally exhibit agreement. When they do so it is mostly by means of a coreferential clitic rather than a verbal affix, as in Bulgarian or Spanish, for instance. An example of the even rarer phenomenon of agreement with an adpositionally marked recipient by means of affixal marking is presented in (9) from Burushaski.

(9) Burushaski (Wilson 1996: 35)
jé-e dasin-mo r hán gitáap-an mu-chí-abayam
I-ERG girl-OBL.F to one book-INDEF.ABS 3SG.F-give-1SG.PRES.PRF
'I have given the girl a book.'

It must be pointed out though that in Burushaski agreement with the R (as opposed to the T) occurs only with four verbs.

Given that the degree of overt double marking of each of the verbal arguments is quite low, it comes as no surprise that overt case and agreement marking of both the A and P is quite exceptional. Only 19 (4%) languages in our sample display such double marking. Overt marking by case and agreement of each of the three arguments in a ditransitive clause does not seem to be attested.[6]

As for the distribution of overt case and agreement marking relative to word order type, recall from section 19.3 that the pattern of marking that is seen to be most advantageous for on-line processing is the case marking of preverbal arguments and the agreement marking of postverbal ones. Among the languages in our sample, the percentage of case-marking-only verb-final languages is higher than that of verb-initial and verb-medial ones (19% vs. 15% vs. 5%), and is statistically significant at the 0.5% level. And among the verb-initial languages, agreement-only ones are indeed more common than those manifesting both agreement and case marking or case marking alone (52% vs. 25% vs. 15%). But the frequency of agreement-only languages among the verb-initial ones is not much higher than among verb-medial ones (45%). Thus in all, while there are differences in the intra-language distribution of overt case and agreement marking relative to word order type, they are not as extensive as might be expected. This is to a large extent attributable to the high frequency of agreement marking in all word order types, which in turn is a reflection of its referent tracking function rather than its argument discriminating one.

[6] Both the T and R may take affixal case marking or the T may be marked by a case affix, the R by an adposition.

978–0–19–920647–6 20-Spencer-ch19 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 303 of 303 May 16, 2008 18:58

19.5 Final remarks

In our discussion of the relationship between case marking and word order and case marking and agreement marking we have concentrated on the extent to which the three forms of encoding may complement each other and the extent to which they interact with each other. We have argued that at least for lexical NPs the discriminatory function of case marking is better rendered by word order than by agreement marking. This is reflected in the existence of a statistically significant correlation between case marking and basic word order type. Significantly, there is no comparable correlation between agreement marking and word order type. We have also shown that the likelihood of an argument displaying both overt case and agreement marking declines as we progress down the argument prominence hierarchy and attributed this distribution to the indexing properties of case marking which are more strongly in evidence in relation to Ps and Rs than As. Finally we have documented that, despite assumptions to the contrary, there are only very weak patterns with respect to the intra-language distribution of case and agreement marking relative to word order type.

An aspect of the interaction between case, agreement, and word order which we did not explore is the association between the presence of the two morphological forms of encoding and word order flexibility. The basic insight underlying this association is that if the A is morphologically distinguished from the P by case and/or agreement marking, word order is 'freed' from performing this function and can thus be used for other communicative ends. While the cross-linguistic studies of Steele (1978) and subsequently Siewierska (1998) have lent some support to the above, especially to the lack of word order flexibility arising from the absence of case and agreement marking, the lack of consensus on which word order variants should be taken into account in determining word order flexibility considerably reduces the efficacy of any concise discussion of the issue. A comparison of the effects on word order flexibility of the presence of case marking with that of the presence of agreement marking therefore deserves more space than can be given to it here.

CHAPTER 20

CASE MARKING AND ALIGNMENT*

BALTHASAR BICKEL
JOHANNA NICHOLS

20.1 Introduction

Alignment is standardly illustrated with ergative vs. accusative coding of subjects as in Basque (1) and Russian (2):[1]

(1) a. *gizona etorr-i da*
man.NOM arrive-PRF.PTCP AUX.3SG.S
'The man has arrived.'
b. *gizona-k mutila ikus-i du*
man-ERG boy.NOM see-PERF.PTCP AUX.3SG.A3SG.O
'The man has seen the boy.'

* Research for this chapter was facilitated by DFG grant No. BI 799-3/1 (Bickel) and NSF grant no. NSF 96-16448 (Nichols).

[1] Throughout this chapter we use *nominative* to refer to the morphological case that is citation form and marks S, regardless of whether S is aligned with A or O (though many linguists use nominative only for S=A and *absolutive* for S=O). As the rest of the chapter shows, there are so many cross-cutting alignment patterns that a terminology treating morphological cases as nodes in an alignment space and separately labelling every node would be cumbersome, as well as confusing morphological case paradigms with alignment patterns.

(2) a. *muzhchina prishel*
man.NOM came.PAST.M
'The man arrived.'
b. *muzhchina uvidel devushku*
man.NOM see.PAST.M girl.ACC
'The man saw the girl.'

In (1) the syntactic functions of S (*gizona* 'man' in 1a) and O (*mutila* 'boy' in 1b) are identically coded, while in (2) it is S (*muzhchina* 'man' in 2a) and A (*muzhchina* in 2b) that are identically coded. That is, the two languages differ in whether S is aligned with A or O. Henceforth we refer to syntactic functions such as A, O, and S as *argument roles.* Argument structure has to do with the number of arguments and their roles; valence has to do with the morphosyntactic treatment of these roles, including their marking with particular cases.

Alignment is the identical vs. distinct coding or treatment or behaviour of argument roles that are different at some other level or in some other part of the grammar. Put differently, alignment is neutralization of valence-specific argument roles in particular morphological or syntactic contexts. Alignment, from our perspective, holds between sets (usually pairs, sometimes triads) of argument roles that have the same formal treatment (e.g. case) in some context in some language, and alignment patterns define construction-specific and language-specific grammatical relations. The typological and theoretical issues in alignment include ascertaining the possible alignment types, extending the theoretical and descriptive apparatus for alignment beyond the standardly covered A, S, and O, determining the contexts and factors that constrain or favor alignment types (such as referential hierarchy effects), and accommodating them in a comprehensive typology of alignment. In this chapter we are concerned with alignment as identified by coding in cases, though morphological alignment is also manifested in verb agreement, and syntactic alignment appears in such things as word order, argument sharing, conjunction reduction, aspects of complementation, accessibility to relativization and the like, pivot in valence-changing derivations, and control of reflexivization and non-finites. Alignment of case marking can be seen as a generalized version of morphological syncretism: the same marker covers different argument roles but, unlike classical syncretism, it does so across all, or nearly all, paradigms.

It is convenient to be able to identify a language as exhibiting some alignment type: Basque is (morphologically) ergative, Russian accusative, Lakhota stative–active (or split-intransitive), Vietnamese neutral. However, this is rarely, and probably never, accurate, for three reasons. First, there are lexical splits: most and probably all languages have verbs that display distinct valence patterns. For instance, Basque has enough agentively inflected subjects, and Russian enough dative experiencer subjects, to qualify either language as split-intransitive. Second, there

are grammatical splits. The alignment of the case marking is not necessarily the same as that of other parts of the grammar: a number of languages, like Georgian, have ergative case marking but accusative verb agreement. Alignment may vary with tense and other factors: again in Georgian, case marking is ergative in the past tense series but accusative in the present series. (This describes a large and open part of the Georgian verbal vocabulary, but Georgian also has a significant lexical split and most specialists describe it as stative–active.) Third, alignment is often taken to pertain only to the coding of S, A, and O, but the relative coding of different objects and the coding of arguments like possessors or like non-arguments are also kinds of alignment.

We define our purview as follows. First, we consider case marking in the broad sense: morphological cases (whether marked by affixes, clitics, separate case words, ablaut, tone changes, etc., all of which we regard as differing in their degrees of phonological fusion but not as different grammatical phenomena) and adpositions (at least in languages where the adpositions govern cases and the same argument role can be marked by a bare case in one context but an adposition plus case in another, as in most Indo-European languages that preserve cases). The same alignment patterns can be found in head-marking systems as well, but we do not cover these. The alignment of syntactic phenomena, which has to do with which arguments behave alike rather than which are marked alike, is also not covered here (but see Bickel 2007a for a recent survey.)

Second, we ultimately attempt to catalogue all alignment types involving arguments, including those where arguments are coded as adjuncts or possessors, but we do not attempt to cover alignment of all adjuncts or all possessors, and for reasons of space, we concentrate here on argument alignment.[2] So approached, alignment types are nearly the same thing as valence types, and we define alignment types as subsets of arguments defined by their roles in valences plus their referential properties. Surveying alignment patterns for any language then requires surveying all the valence types found in that language, or at least those found with any frequency. Alternatively, one exhaustively surveys all the predicate-specific functions of each individual case (e.g. dative marks S of lexical set 1, A of lexical set 2, G of all verbs except lexical set 6, etc.).

20.2 Argument roles

Predicates license a specific number of arguments: zero (*snow*), one (*run*), two (*see*), three (*give*). Arguments are therefore identified first of all by numerical valence.

[2] An expanded version of this chapter that deals systematically with alignments of arguments and possessors will appear elsewhere (Bickel and Nichols, forthcoming).

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 307 of 321 May 16, 2008 18:58

This distinguishes S, the sole argument of one-place predicates, from all others. The arguments in two-place or three-place predicates can be distinguished from each other by differences in their semantic entailments (see Dowty 1991, Primus 2006a, among others):

A1 The more agent-like argument of a two-place predicate
O The less agent-like argument of a two-place predicate
A2 The more agent-like argument of a three-place predicate
G The more goal-like non-agent-like argument of a three-place predicate
T The non-goal-like and non-agent-like argument of a three-place predicate

An argument is more agent-like if it causes an event, if it is volitional or at least sentient, and/or if it exists independently of the event. Thus, in an experiential predicate, experiencers always qualify as more agent-like than stimuli. Since agent-like properties also characterize the most common topics in discourse, measures of topicality are a useful rule of thumb for evaluating whether an argument is agent-like or not. (The more topical an argument is, the more likely it is agent-like.) An argument is more goal-like if it is stationary in movements, if it is the target of a communicative event, or if it benefits or suffers from an event.

A1 and A2 are most often aligned with each other, and we use 'A' as a cover symbol neutralizing the distinction. The only language with a general distinction between A1 and A2 in case marking that we know of is Gyarong:

(3) Gyarong (lCog-rtse rGyal-roṅ) (Sino-Tibetan; Sichuan; Nagano 1984)
a. *nəyo-ki chigyo kəw-nasṅo-ch ko*
2S-ERG 1DU.NOM 2>1-scold-1DU AUX
A1 O
'You scold us.'
b. *nəyo chigyo kəw-wu-ch ko*
2SG.NOM 1DU.NOM 2>1-give-1DU AUX
A2 G
'You give (it to) us.'

(3a) is monotransitive, and its A1 argument is obligatorily marked by the ergative case in *-ki*, while the A2 argument of (3b) cannot take the ergative case.[3] In addition, several languages have a covert difference in that A1 exhibits minor alignments such as A1=G (e.g. with experiencers marked dative, as in (9) and (10) below), while A2 apparently never does. This is probably because A2, in all instances known to us, is uniformly fairly agent-like, unlike A1, which ranges over various semantic roles (agent, experiencer, recipient, location). Put differently, ditransitives allow

[3] As important as this observation is for appreciating the true range of typological variation, the distinction between A1 and A2 is currently becoming lost among younger speakers of Gyarong, probably because of increased exposure to Chinese (Y. Nagano, personal communication, October 2003).

only agentive A's, while monotransitives allow agentive, experiencer, etc. A's. This means that A1 and A2 are different generalized roles, supporting our decision to regard them as different argument roles.

Apart from these argument roles, alignment sometimes extends to adjuncts (abbreviated here as Ad) and to adnominal dependents (abbreviated as Poss, for 'possessor', the most common semantic role of adnominals).

Argument roles are defined here by the minimal distinctions necessary in numerical valences. Individual predicates make much more fine-grained distinctions (known as semantic or thematic roles), and in many languages lexical and/or semantic distinctions between such predicates condition case alignment: e.g. one-place predicates in Basque are divided into those that take agents and others (a pattern known as split intransitivity). Those that take agents align S with A, while those that take non-agents align S with O. A very common lexical split of this kind involves two-place predicates where the O argument represents a goal. Such predicates often fail to follow the general alignment pattern of O arguments and instead align these arguments with adjuncts. An example is English *go to X*. (That X is still an O argument is evident from the fact that *go* assigns a goal role even in the absence of the preposition *to*: in *Where did she go?, where* must be interpreted as a goal, in contrast to *Where did she walk*?, in which *where* can be either a location or a goal.) We return to this issue in section 20.3.6 below.

20.3 Alignment possibilities

This section illustrates the alignment patterns we are aware of. Probably every imaginable alignment pattern is found somewhere in some language, but the ones of chief typological interest are those that are the major or dominant or basic type in one or another language and those that are fairly frequent or the obvious major contenders in splits. Traditional surveys suggest that these are alignments of S with A ('accusative') or O ('ergative'), and of O with either G ('primary object'; 'secundative' in Haspelmath's 2005a terms) or T ('direct object' or 'indirective'). We refer to alignments including S (S=A, S=O, S=A=O, etc.) as 'S-alignments' and to alignments including O (O=G, O=T, O=G=T, etc.) as 'O-alignments'.

For expository reasons, we typologize alignments primarily by the way S is aligned, and, within these types, by the way objects are aligned. In reality, S-alignment and O-alignment are logically independent variables, and only a large-scale quantitative survey can establish distributional preferences and clusters.

Examples below have an additional interlinear line showing their argument roles. For each example we indicate whether this alignment pattern is a major type, important though not major, or a minor type in that language.

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 309 of 321 May 16, 2008 18:58

20.3.1 Tripartite-based alignments

Full-fledged tripartite systems would have distinct cases for all major roles, i.e. S≠A≠O≠T≠G. However, in all cases of languages with S≠A≠O that we are aware of, at least one pair of the object roles is aligned, either T with O or G with O.

In Yazgulyam, objects align following a direct object pattern, with T=O. The resulting S≠A≠O=T≠G pattern is limited to pronouns. (Nouns only have dative marking on G and some prepositionally governed arguments.)

(4) Yazgulyam (Iranian; Pamir; Èdel'man 1966)

a. *áz=əm mɔt mad*
1sABS=1s tired become.PST
S
'I am tired.' (Èdel'man 1966: 37)

b. *mon š-tu wint*
1SG.OBL ACC-2SG.OBL see-PAST
A1 O
'I saw you.' (Payne 1980: 176)

c. *ž-way aǰa ni waðók-ra*
ACC-3SG.M.OBL give.IMP for puppet-DAT
T Ad
'Give this for my puppet!' (Èdel'man 1966: 167)

d. *aǰa mó-ra k'aðit*
give 1SOBL-DAT key
G T
'Give me the key!' (Èdel'man 1966: 175)

In Nepali, O aligns with G, but in addition, T also aligns with S. The resulting pattern is T=S≠A≠O=G and is limited to past tense sentences with definite O arguments; when these conditions are not met, alignment is accusative and based on direct objects (see Bickel 2007a):

(5) Nepali (Indo-European; Nepal)

a. *ma ga-ẽ*
1SG.NOM go-1SG.PST
S
'I went.'

b. *mai-le Rām-lāi dekh-ẽ*
1SG.-ERG Ram-DAT see-1SG.PAST
A1 O
'I saw Ram.'

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 310 of 321 May 16, 2008 18:58

c. *mai-le* *kitāp* *Rām-lāi* *di-ē.*
1s-ERG book.NOM Ram-DAT give-1SG.PAST
A2 T G
'I gave the book to Ram.'

20.3.2 S=A alignments

Accusative alignment is found with all kinds of O-alignment patterns. Ute illustrates S=A≠O=T=G[4] alignment; the object case is distinguished from the nominative by voiced vs. voiceless finals:

(6) Ute (Uto-Aztecan; Colorado; Givón 1980)
a. *tá'waci̯* *wu̯u̯ka̯-xa*
man.NOM work-PAST
S
'The man worked.'
b. *mamáci̯* *ta'wáci* *pu̯níkaay-kya*
woman.NOM man.OBJ see-PAST
A1 O
'The woman saw the/a man.'
c. *mamáci̯* *ta'wáci* *pu̯sáriniyu̯pi* *máy-kya.*
woman.NOM man.OBJ story.OBJ tell-PAST
A2 G T
'The woman told the man a story.'

When accusative alignment combines with primary objects (i.e. S=A≠O=G≠T), the T argument is typically coded like the S and A argument, i.e. there is no distinct T case and the actual pattern is S=A=T≠O=G. A frequent further restriction is that such patterns are mostly limited to O arguments that rank high on the referential hierarchy, following principles of differential object marking. Low-ranking O arguments then generally appear in the (unmarked) nominative case like S and A. We exemplify the pattern with Awa Pit; other examples would be Spanish, Persian, and many other languages. (7a) shows (zero-marked) nominative case on S, and the same case is found on all A arguments in (7b–d). Human and other higher-ranking O arguments align with G and are marked by dative clitic *=ta*, as in (7b,c). Note that the T argument of ditransitive is always in the nominative, whether it is human or not.

[4] A formula with an internal '≠' abbreviates a sequence of alignment statements: S=A; A≠O; O=T=G.

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 311 of 321 May 16, 2008 18:58

(7) Awa Pit (Barbacoan; Colombia; Curnow 1997)

a. *na=na* *nayŋ-ma-ṫi-s*
1SG.NOM=TOP fall-COMPL-PAST-CONJUNCT.UNDERGOER
S
'I fell.'

b. *kin-ka=na,* *na=na* *Santos=ta*
dawn-when=TOP 1SG.NOM=TOP S.=DAT
A1 O
izh-ta-w
see-PAST-CONJUNCT.UNDERGOER
'At dawn I saw Santos.'

c. *Camilo=na* *Santos=ta* *pala* *kwin-ṫi-zi*
C.NOM=TOP S.=DAT plantain.NOM give-PAST-DISJUNCT
A2 G T
'Camilo gave Santos a plantain.'

d. *na=na* *Santos=ta* *pashu* *mi̇la-ta-w*
1SG.NOM=TOP S.=DAT daughter.NOM give-PAST-CONJUNCT.SBJ
A2 G T
'I gave my daughter to Santos.'

Direct object patterns combine with accusative alignment more freely. A very common pattern is the one illustrated here by Khasi: S=A≠G≠O=T

(8) Khasi (Austroasiatic; India; Nagaraja 1985)[5]

a. *u-briew* *u-la-wan*
3sm-man.NOM 3sm-PAST-come
S
'The man came.'

b. *u-briew* *u-la-pi̇nyap* *ya-u-bseñ*
3SG.M-man.NOM 3SG.M-PAST-kill ACC-3SG.M-snake
A1 O
'The man killed a snake.'

c. *u-khi̇nna?* *u-la-a꞉y* *ya-u-khulam* *ha-ka-koŋ*
3SG.M-boy.NOM 3SG.M-PAST-give ACC-3SG.M-pen DAT-3SG.F-sister
A2 T G
joŋ-u
GEN-3SG.M
'The boy gave a pen to his sister.'

[5] We write some preposed morphemes as prefixes here because they are grammatically bound to stems and cannot occur independently. But they are phonologically separate words, and this is why they are typically not written as prefixes but as preposed words.

With primary objects and accusative alignment we noted that the ditransitive object that is excluded from the object case (here, T) tends to align with S and A; with direct objects, where the excluded argument is G, this is less common, but it does occur as a minority pattern: S=A=G≠O≠T. This is not uncommon for experiential predicates and can be exemplified by Russian, where it represents a minority but fairly important pattern:[6]

(9) Russian
a. *Mne* *strashno.*
1SG.DAT scary
S
'I'm scared.'
b. *Bol'she* *vsego* *mne* *tam* *ponravilas'* *pogoda.*
most of.all 1SG.DAT there like.PAST.F weather(F)-NOM
A1 O
'I liked the weather most of all.'
c. *Daj* *mne* *kuklu*
give-IMP 1SG.DAT doll-ACC
G T
'Give me the doll.'

S=A=G patterns are also often found as minority patterns among experiential predicates in languages that strictly differentiate S and A with most other predicates. An example is Ingush, where most other predicates show ergative alignment:

(10) Ingush (Nakh-Daghestanian; Caucasus)
a. *Suona* *shiila* *jy*
1SG.DAT cold J.is
S
'I'm cold'
b. *Suona* *yz* *bwarjg+veira*
1SG.DAT 3SG.ABS(V) eye+V.see.WP
A1 O
'I saw him'
c. *Aaz* *cynna* *mashen* *j.elar*
1SG.ERG 3SG.DAT car(J)[NOM] J-gave
A2 G T
'I gave him/her a car'

[6] Based on the semantic entailments (cf. section 20.2), we regard the dative experiencer in these examples as A, though this is not settled. There is a large literature on the question; see e.g. Barðdal 2004, Bickel 2004a, Bossong 1998, Marušič and Žaucer in press, Perlmutter and Moore 2002, Nichols 2007, forthcoming.

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 313 of 321 May 16, 2008 18:58

The mirror image of this is S=A=T≠O≠G and is found with accusative marking on experiencers:

(11) German
a. *Ihn friert.*
3SG.ACC be.cold.3SG.NPST
S
'He is cold.'
b. *Ihn interessiert das nicht.*
3SG.ACC be.interesting DEM.NOM not
A1 O
'He doesn't find this interesting.'
c. *Sie übergaben ihn der Polizei.*
3SG.NOM hand.over 3SG.ACC ART.SG.DAT police
A2 T G
'They handed him over to the police.'

This analysis follows our methodological principle of analysing argument roles exclusively by their semantic entailments (section 20.2). The A and S arguments here do not share properties with other A arguments beyond the semantics, showing that other constructions – specifically verb agreement – have another alignment in German (S=A≠T=G=O throughout).[7]

20.3.3 S=O alignments

Ergative patterns with neutral O-alignment (A≠S=O=T=G) is illustrated by Belhare, where it is the majority pattern in the lexicon:

(12) Belhare (Sino-Tibetan; Nepal)
a. *a-phu ta-he*
1SG.POSS-eB[NOM] [3SG.S]come-PAST
S
'My elder brother came.'
b. *a-phu-ŋa a-tak nis-e*
1SG.POSS-eB-ERG 1SG.POSS-friend[NOM] [3SG.A]see-PAST[3SG.O]
A1 O
'My elder brother saw my friend.'

[7] Note that third person singular agreement in *Mich friert* is a default form and not triggered by the S argument. Although the sentence can be expanded into *mich friert es*, nominative *es* 'it' does not realize an argument, as can be seen from the fact that it cannot be replaced by a lexical NP (**das Wetter friert mich* 'the weather is cold to me').

c. *un-na* *a-tak* *celi*
3SG.-ERG 1SG.POSS-friend[NOM] clan.sister[NOM]
A2 G T
pir-he
[3SG.A]give-PAST[3SG.O]
'He gave a *celi* (marriageable agnatic relative) to my friend.'

The combination of ergative and primary O-alignment, i.e. A≠S=O=G≠T is shown by Inuit (West Greenlandic), where it is the standard, majority pattern:

(13) Inuit (Eskimo-Aleut; Greenland; Manning 1996)
a. *Oli* *sinip-p-o-q.*
O.-NOM sleep-IND-INTR-3SG
S
'Oli sleeps.'
b. *Aani-p* *miiqqa-t* *tama-isa* *taku-nngi-la-i*
A.-ERG child-PL.NOM all-3PL see-NEG-IND-3SG>3PL
A1 O
'Aani saw none of the children.' / 'Aani didn't see all of the children.'
c. *Juuna-p* *miiqqa-t* *atuakka-mik* *nassip-p-a-i*
J.-ERG child-PL.NOM book-MODAUS send-IND-TR-3SG>3PL
A2 G T
'Juuna sent the children a book.'

However, ergative (non-tripartite) alignment seems to more commonly align with direct objects. The pattern A≠S=O=T≠G is a major pattern in Ingush (involving productive derivational machinery):

(14) Ingush (Nakh-Daghestanian, Caucasus)
a. *Marem* *qiera-jalar*
M.NOM fear-J.INCHOATIVE-WP
S
'Mariam got scared'
b. *aaz* *Marem* *qiera-jyr*
1SG.ERG M.NOM fear-J.CAUS-WP
A1 O
'I frightened Mariam'
c. *Aaz* *cynna* *mashen* *j.elar*
1SG.ERG 3SG.DAT car(J).NOM J-gave
A2 G T
'I gave him/her a car'

As noted above, an important minority of verbs in the Ingush lexicon aligns A=S=G≠O=T.

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 315 of 321 May 16, 2008 18:58

20.3.4 A=O alignments

Especially among experiential predicates it is common for A arguments to be coded as O, G, or T arguments, but, as we noted in the discussion of S=A types above, this usually (although to different degrees) extends to experiential S arguments as well. A limiting case is Latin, where accusatively marked experiencers are virtually limited to two-place predicates, that is, there are almost no accusative-marked S arguments (unlike in German, cf. example (11a) above):

(15) Latin

ut	*me*	*non*	*solum*	*pigeat*	*stultitae*	*meae,*
COMP	1SG.ACC	not	only	feel.bad3SG.SBJV.PRS	stupidity.GEN	my.GEN
	A1				O	

sed etiam pudeat (Cicero, *De domo sua* 29)
but also be.ashamed.3SG.SBJV.PRS
'so that I not only feel bad about my stupidity but that I'm also ashamed of it'

However, note that in examples like these the O argument (*stultitae meae* 'my.GEN stupidity.GEN') aligns with possessors and therefore does not show the same case as the A.

Apart from such minority patterns, it is rare for A arguments to align with objects but not also S. This does occur, however, as the majority pattern in a few Iranian languages of the Pamirs, e.g. Rushan, where it is limited to the past tense:

(16) Rushan (Indo-European: Iranian; Afghanistan; Fajzov 1966: 57, 202, 61)

a. *az=um* *tuyd*
1SG.NOM=1S go-PAST
S
'I went'

b. *mu* *way* *wunt*
1SG.OBL 3SG.OBL saw
A1 O
'I saw him.'

c. *tā* *mu* *kā* *talēpt?*
2SG.OBL 1SG.OBL why sought
A1 O
'Why did you look for me?'

Here both A and O, but not S, appear in the oblique case. The nominative is reserved for S arguments.

20.3.5 Neutral alignment

Neutral S-alignment combines both with primary objects and direct objects (and of course with neutral O-alignment, which is the same as the absence of argument case marking). Ju|'hoan illustrates the primary object pattern, i.e. S=A=O=G≠T:

(17) Ju|'hoan (Ju; Angola, Botswana, Namibia; Dickins n.d.)

a. *jú* *n!ànì* *tsí*
people three come
S
'Three people have come.'

b. *dà'ábí* *|óá* *ho* *n!àmà.*
children NEG find road
A1 O
'The children did not find the road.'

c. *dà'ámá* *jàn* *|'àn* *ha* *bá* *kò* *màrì.*
child good give 3SG father OBL money
A2 G T
'The good child gave his father money.'

Neutral alignment combined with a direct object pattern, i.e. S=A=O=T≠G, can be illustrated with French or English, where a preposition (*à*, *to*) marks the G argument, while all other (non-pronominal) arguments are treated the same way and have no overt case marker.

Neutral alignment most often involves zero morphological exponence. When there are neutrally aligned overt case markers, their use is often governed by the ranking of arguments on the referential hierarchy. The proximative (also known as nominative) case in Tagalog, for example, is assigned to the argument that is most topical in discourse (identified here as such by italics in the translation):

(18) Tagalog (Austronesian; Philippines, Kroeger 1993)

a. *bumili* *ang=lalake* *ng=isda* *sa=tindahan.*
PFV.A.buy PROX=man OBL=fish LOC=store
A O Ad
'*The man* bought fish at the/a store.'

b. *binili* *ng=lalake* *ang=isda* *sa=tindahan.*
PFV.O.buy OBL=man PROX=fish LOC=store
A O Ad
'The/a man bought *the fish* at the/a store.'

c. *binilhan* *ng=lalake* *ng=isda* *ang=tindahan*
PFV.G.buy OBL=man OBL=fish PROX=store
A O Ad
'The/a man bought fish *at the store.*'

The other core case of Tagalog, marked by *ng*, also has neutral alignment but is reserved for low-ranking arguments. Similar systems are found in a number of languages of the Americas (see Zúñiga 2006 for a survey) and are commonly refered to as 'hierarchical alignment'. It should be noted, however, that referential ranking also plays a central role in non-neutral kinds of alignment. Referential ranking, in our view, is not an alignment but a secondary, referentiality-based and often discourse-related, elaboration of a basic alignment (e.g. neutral in the case of Tagalog).

20.3.6 Adjunct and possessor alignment

All examples discussed so far have arguments coded the same as other arguments. But arguments also often align with adjuncts or possessors in the way they are treated by case or other constructions. The most well-known instances of adjunct alignment are those of languages where A arguments receive the same case marker as instrumentals or ablatives. This phenomenon is often referred to not as alignment but as syncretism (ergative–instrumental and ergative–ablative syncretism, respectively). If the pattern is general across all paradigms of a language, it is different from syncretism, in the same way as the use of the same case marker for S and O across all paradigms of a language is different from syncretism.

Alignment of O arguments with adjuncts is also frequent – for example in the form of adpositional object marking such as the use of the preposition *a* before animate or specific objects and for locational adjuncts in Spanish; or in the form of the multitude of adpositions used on specific verbs in Russian.

Alignment of arguments with possessors is best known in the form of case markers covering both genitive and ergative functions, as e.g. in Eskimo languages, or in the form of genitives on experiential and S or A arguments, as in the following example from Bangla:

(19) Bangla (Indo-European; India and Bangladesh; Klaiman 1980)

a. *āmār* *āscorjo* *ho-lo.*
1SG.GEN surprise become-3PAST
S
'I was surprised.'

b. *āmār* *tomāke* *cāi.*
1SG.GEN 2SG.DAT need.3
A1 O
'I need you.'

Another well-known example is genitive-marking on O arguments in Latin and several other European languages (see Haspelmath and Michaelis forthcoming for a recent survey.)

Further discussion of adjunct and possessor alignment can be found in Bickel and Nichols (forthcoming).

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 318 of 321 May 16, 2008 18:58

20.4 The stative–active type

Split-S, or stative–active (or similar terms), is generally taken to be an alignment type on a par with accusative, ergative, etc. Languages with this alignment type are considered to have no single basic subject alignment, and the coding of S is based on agency, Aktionsart, or a similar factor. There are several respects in which split-S is not a satisfactory alignment type, however. First, nearly every language has at least some verbs with oblique or otherwise atypically coded S; the difference between split types and unsplit types is one of degree. Second, as Sapir (1917) first noted and Merlan (1985) first showed in detail, in some split-S languages the verbs with O-coded S are a closed, small, or otherwise delimited class and A-coded S verbs are open and productive; while in others it is the A-coded set that is closed, small, or delimited while O-coded S is open and productive. This makes it possible to consider such languages respectively accusative and ergative, with a larger than usual set of oblique-S verbs but still a clear basic alignment type. (There are some languages in which both sets of verbs are large enough that it is difficult to discern a basic type; Georgian is one such. There are, of course, fluid-S languages like Acehnese or Batsbi in which nearly every verb can take both S coding types depending on semantic factors; fluid-S languages are not at issue here. Note, however, that Batsbi is fluid-S only in the first and second persons; the third person is ergative.) Third, not only the sizes but the lexical membership of the classes differ from language to language, often unpredictably. Fourth, languages fall into a continuous cline running from ergative to accusative via split-S, with no cut-off points that might define discrete types and considerable overlap between languages that are generally considered stative–active and ones that are not. (Nichols forthcoming surveys twenty verb glosses across forty-one languages, and when the number of O-coding verbs is plotted against the number of A-coding verbs a continuous cline results.) Fifth, split-S languages generally have the same split in A: two-argument verbs like 'like', 'forget', and 'remember' are quite likely to have O-coded A. Sixth, S=O is too general a description of the object-coded type; the great majority of split-S languages have O=G alignment, so the essence of the type is S=G (or S=O=G) coding, and this means that the dative-coded experiencers of many Eurasian languages (which are O=T) exhibit the same pattern. Seventh, if one posits split-S as a separate type, one would, by the same logic, also need to posit split-O types, where again the split falls into distinct patterns following O=T, O=G, or O=Ad alignments, and ultimately also split-G and split-T types.

These problems are inherent in seeking discrete grammatical types when the grammatical patterning is driven by individual lexical items. Probably the best way to define types is to set up a standard list of verb glosses, determine the argument coding of each of those verbs in the language in question, and use the frequencies of the different coding types and their distribution in the list to typologize languages.

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 319 of 321 May 16, 2008 18:58

20.5 Compound verbs

A number of languages, including most languages of southwest Asia, have periphrastic compounding with light verbs as a main means of verbal derivation. Often the heavy element of the compound is nominal in origin and retains some of the grammatical properties of independent nouns:

(20) Persian
man nafas mi-kesh-am
1SG breath CONT-pull.PRES-1SG
S
'I breathe'

(21) Ingush
aaz sa doax
1SG.ERG breath(D) D.take:PLURACTIONAL
S
'I breathe'

In these examples, as is common, the light verb is a transitive verb that takes the nominal element as its direct object. We regard this as the etymologically interesting but syntactically irrelevant internal morphological structure of the verb, and on this analysis the subject is an S and the verb has one argument. As a consequence, a number of derived verbs in an ergative language like Ingush have ergative-marked S arguments, i.e. S=A alignment.

A similar issue arises with the light verb constructions common in Southeast Asian and Himalayan languages, called *psycho-collocations* by Matisoff (1986). Typically these consist of a body-part or experiential noun as heavy piece, with the experiencer coded as its possessor; compare the examples from Bangla in (19)). The resulting structure is again one where the experiencer is syntactically an S (or A) argument although morphologically and etymologically it is a possessor.

Exceptional O=Poss alignments, as with genitive-marked objects in Latin, also tend to have their roots in compound constructions where the genitive used to be in an adnominal relation to a lexical noun.

20.6 Non-lexical conditions

The choice of case alignments is most often lexically determined by the verb. But other factors condition alignment as well. Tense and aspect categories commonly

978–0–19–920647–6 21-Spencer-ch20 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 320 of 321 May 16, 2008 18:58

condition alignment in Eurasian languages with tense-based split ergativity. Often the explanation of such splits involves a mix of pragmatic factors and the specific nature of the morphological forms involved; for a case study on the history of tense-conditioned alignment splits in Indo-Aryan, see Peterson (1998).

The referential hierarchy is associated with a split where high-referentiality NP's (personal pronouns, human nouns) have accusative or at least non-ergative inflection while lower-referentiality ones have ergative inflection. (See Bickel 2007a, forthcoming, a, for some recent discussion).

Alignments may be split between main and subordinate clauses, in nominalized vs. verbal predicates, etc. Dixon 1994: 97–104 lists languages with some of these kinds of splits.

20.7 Whole-language types?

As noted in the introduction, it is standard to pigeonhole languages as accusative, ergative, stative–active, etc. and as primary/secondary or direct/indirect object. As almost all our examples above show, this is not correct in detail; every language we have seen has at least some verbs exhibiting other patterns, at least some of which are in fact salient in the language, and some languages have two clear contenders for default or plurality type. In addition, the typological literature on alignment has had very little to say about oblique coding of arguments and about compound and light verbs, both of which result in additional minor and sometimes even major splits in alignment.

But for many research questions it is still of great interest to be able to classify languages (or even whole families) by their alignment type. As with 'stative–active' splits, the proper way to do this is to count lexical frequencies of alignment patterns on some controlled wordlist and also text frequencies on a sufficiently large and varied corpus of authentic text to give a reliable reflection of the language's type. The only attempt in this direction that we know of is Nichols (forthcoming), a pilot lexical study on split S and A coding.

An alternative approach is to survey all distinct alignment patterns in each language and derive an aggregate value estimating the overall trend in the language. For example, of all case alignment patterns in Russian, two include an alignment of S with A (nominative and dative) and an appropriate aggregate value would qualify this language as being more S=A-like than, say, Ingush, where only one pattern (the one with datives) aligns S with A. Obviously, it is not enough to determine whether S aligns with O; it is also important to know whether O, in turn, aligns with T or G; and ideally also which argument, if any, aligns with adjuncts or possessors.

Adequate quantitative surveys of alignment patterns ultimately need both approaches, with aggregate values based on exhaustive listings of available alignments in each language, weighted for lexical and discourse frequency.

20.8 Conclusions

A typology of alignment must necessarily be relativized to lexical valence sets and must ultimately include the alignment of all argument roles among themselves and with all sorts of adjuncts and nominal dependents. In this chapter we have concentrated on the better known alignments of arguments with arguments.

We have not discussed the distribution of alignment types in this chapter. For some recent discussion of the effects of the referential hierarchy on case distributions, see Bickel (2007a, forthcoming, a). The basic finding of these studies is that despite their popularity (starting with Silverstein 1976 and Comrie 1981a and continued by Dixon 1994 and Aissen 2003), such effects cannot be considered well-established universals (comparable to, say, the correlation of VP and PP order), and much more empirical groundwork is needed before any such claims can be accepted as proven.

CHAPTER 21

CASE AND VOICE

CASE IN DERIVED CONSTRUCTIONS

MASAYOSHI SHIBATANI

Voice is understood here as the pattern of the form–function correlation along the parameters pertaining to the evolutionary properties of an action (see Shibatani 2006). There are thus marked voice categories pertaining to the origin of an action, e.g. the nature of the agent (spontaneous, passive, causative); those pertaining to the nature of the development of an action, e.g. the affectedness of the patient (middle, antipassive); as well as those pertaining to the termination of an action, e.g. the affectedness of other entities than the patient (applicatives, external possession). Besides these conceptual dimensions, voice phenomena are also controlled by the pragmatic factor of discourse relevance (e.g. inherent and contextual discourse topicality) of the event participants. While the conceptual underpinnings are clearer in some voice categories (e.g. the spontaneous, the middle) than others (e.g. the passive), where the pragmatic motivation may prevail, the conceptual and pragmatic factors converge on grammatical phenomena including voice opposition because they are manifestations of the central communicative principle of relevance in the two functional domains of grammar, namely its conceptual basis and its use in context. The notion of transitivity in grammar and discourse is integral to the study of voice. Indeed, it is a theory of voice that provides 'a superordinate semantic notion which will include all the Transitivity components' (Hopper and Thompson 1980: 279).

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 323 of 338 May 16, 2008 18:58

Since the evolutionary properties of an action are determined by the nature of both the participants and the action itself, voice marking can be theoretically realized either in nominal arguments representing the participants or in a verbal form representing the action. The preponderance for verbal voice marking is as an expression of iconicity; that is, a verb representing an action takes on voice marking because voice has to do with the way an action evolves. Below we will see some cases where voice is marked in nominal arguments, i.e. where voice alternations are signalled by the difference in case form. Three typological patterns of voice marking are, thus, observable: (i) verbal marking plus nominal case marking, (ii) case marking alone, and (iii) verbal marking alone. Type (i) seems most prevalent, where the difference in case marking in opposing voice constructions is a consequence of changes in grammatical relations of the nominal arguments, which are in turn due to the structural change signalled by verbal marking. Voice alternation by means of case marking alone – Type (ii) – is less frequently reported, but this may be due to the historical bias on the part of the grammarians, who have tended to identify grammatical categories in terms of (verbal) morphology rather than function. Indicating voice contrast by verbal form alone – Type (iii) – seems to be less favoured compared to the first two types, but this also seems possible in some voice categories, as shown below.

Consideration of space prevents us from going into any detail of causatives, where many interesting case phenomena obtain (Comrie 1976b, Dixon 2000, Shibatani 2002), the middle voice (Kemmer 1993), external possession (Shibatani 1994, Payne and Barshi 1999), inverse and split-ergativity phenomena, all of which are treated in the comprehensive voice framework proposed in Shibatani (2006).

The unmarked category of **active voice** refers to the correspondence pattern of a transitive situation type and a syntactic transitive structure. The prototypical transitive situation is the one in which an action originates in a volitional agent, extends beyond the agent's personal sphere, and terminates in a distinct patient achieving an intended effect on it. The active voice construction paired with the transitive situation type is syntactically transitive having the agent (A) and the patient (P) encoded as two primary grammatical relations – the Subject and Object in nominative–accusative languages, and the Absolutive and Ergative in absolutive–ergative languages. Against the background of this basic, active-voice pairing of the transitive situation and the syntactic transitive structure, opposing marked voice constructions represent deviations from the prototypical transitive situation.

21.1 Spontaneous voice

The spontaneous voice represents a situation where an action is brought about by a non-volitional actor. Spontaneous constructions typically involve a verb form

different from the one found in the corresponding volitional (active) constructions. Because the volitional/spontaneous opposition centres on the status of the initiating agent, whose intentionality has the scope over the entire course of the action including the planning of the intended act and the intended effect upon a patient, we expect (again from the perspective of iconicity) that the verb as well as the nominals representing the non-volitional actor and the unintentionally affected patient can be formally different from those in the volitional voice form. Indeed, this is normally the case, as shown below:

(1) Sinhala (Gair 1990: 17)
   a. *mamə ee wacəne kiwwa.*
      I.NOM that word say.PAST
      'I said that word.'
   b. *maʈə ee wacəne kiyəwuna.*
      I.DAT that word say.P.PAST
      'I blurted that word out.'

(2) Diyari (Austin 1981: 154)
   a. *ŋaṯu yinaɳa ḍanka-ṇa wara-yi*
      1SG.ERG 2SG.O find-PTCP AUX-PRES
      'I found you.' (after deliberately searching)
   b. *ŋaɳi ḍanka-ṯadi-ṇa wara-yi yiŋkaŋgu*
      1SG.ABS find-SPON-PTCP AUX-PRES 2SG.LOC
      'I found you.' (accidentally)

The Sinhala example above illustrates the pattern where so-called dative subject constructions are recruited for the expression of the spontaneous voice. Dative subject constructions in general contain verbal forms which express uncontrolled states of affairs such as possession/existence, ability, psychological and physiological conditions, and as such they are an ideal construction to be recruited for the spontaneous voice. When a language uses a case other than the dative for the non-controlling protagonist in such constructions, the same case is used for the non-volitional actor in the spontaneous voice, e.g. Hindi instrumental *se*, Bengali genitive *-r*, Nepali ablative *-baata*.

To a great extent the case forms of the non-volitional actor and the patient in the spontaneous voice are determined by the structures in which they occur. The typical spontaneous constructions are intransitive, where the non-volitional actor receives an oblique case form, whereas the patient assumes the case form of the sole argument (S) of an intransitive clause in the language.

The two spontaneous constructions illustrated above involve verbal voice marking as well as change in case form of the nominals representing the event participants. The spontaneous expressions, however, can be realized by case marking alone, as shown in the following Lezgian examples, where the contrasting pattern

obtains most systematically with so-called labile verbs, that is, verbs having the same form for transitive and intransitive use:

(3) Lezgian (Haspelmath 1993a: 292)
a. *Zamira.di get'e xa-na.*
Zamira (ERG) pot break-AOR
'Zamira broke the pot.'
b. *Zamira.di-way get'e xa-na.*
Zamira-ADEL pot break-AOR
'Zamira broke the pot accidentally/involuntarily.'

The most celebrated case where case marking alone indicates the volitional/spontaneous contrast is the so-called split-intransitivity phenomenon, as exemplified by the widely cited examples below, where the volitional actor has the agentive form and the non-volitional one the patientive form of a transitive clause.

(4) Eastern Pomo (McLendon 1978: 4)
a. *ha: c'e:xelka*
1SG.A slip
'I am sliding.'
b. *wí c'e:xelka*
1SG.P slip
'I am slipping.'

(5) Tsova-Tush (Batsbi) (Holisky 1987)
a. *(as) vuiž-n-as*
1SG.ERG fall-AOR-1SG.ERG
'I fell down, on purpose.'
b. *(so) vož-en-so*
1SG.NOM fell-AOR-1SG.NOM
'I fell down, by accident.'

The complex evidentiality system of Tibetan languages, on the other hand, allows a rarer form of the volitional/spontaneous contrast, where the opposition is marked solely by verb morphology, that is, by the different evidential auxiliary forms.

(6) Tibetan (Denwood 1999: 137)
a. *ngas. yi.ge. klog.ba yin.*
I-SMP letter read.LINK-AUX (self-centred)
'I read the letter (on purpose).'
b. *ngas. yi.ge. klog.song.*
1-SMP letter read-AUX (other-centered)
'I read the letter (without meaning to).'

While it is arguably the case that the non-volitional actor is conceptually less involved in the event denoted by the verb than the agent of the familiar type of

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 326 of 338 May 16, 2008 18:58

passive, the former is syntactically more integrated to the clause structure than the latter. Spontaneous actors are typically overtly expressed, whereas passive agents are often not syntactically encoded or are even unencodable in many languages. The former, even if encoded in the dative or other oblique cases as in dative subject constructions, participate in syntactic phenomena controlled by the grammatical subject (see the papers in Aikhenvald, Dixon, and Onishi 2001 and Bhaskararao and Subbarao 2004). This syntactic difference between the non-volitional actor of the spontaneous construction and the passive agent reflects greater discourse relevance of the former over the latter (see below). The difference in discourse status and its syntactic reflex between the two types of agents notwithstanding, the conceptual affinity between the two is clearly seen in those languages in which the presence of a non-volitional agent triggers passivization, as in a number of Indonesian languages.

(7) Bahasa Indonesia
   a. *Ali memukul anak itu.*
      Ali AF.hit child the
      'Ali hit the child.'
   b. *Anak itu ter-pukul oleh Ali.*
      child the PASS-hit PREP Ali
      'The child was (accidentally) hit by Ali.'

21.2 Passive and antipassive

The passive and the antipassive have both conceptual basis and pragmatic motivation. The Indonesian accidental passive above highlights a dimension of the conceptual contrast between the active and the passive voice, where the latter conveys a situation in which the agent is a less involved participant in the sense that it is not the protagonist who planned and executed the event with a specific goal in mind. The Mayan language Mam also has a similar accidental passive (see England 1988). English passives such as *John was killed in a traffic accident* and *John was wounded in the war in Iraq* present a case where the involvement of an indefinite agent is viewed as incidental – the emphasis is more on the circumstances under which the named events took place. Conceptually the antipassive correlates with the low degree of affectedness of the patient, which may escape the intended action, as in (10) below, or which may be partially affected, as in (11) below. Passive expressions in which the agent is overtly encoded (e.g. *The letter was signed by the deputy director*) or understood from the context (e.g. *The burglar was finally arrested*) have the pragmatic motivation of downgrading the discourse relevance of the agentive

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 327 of 338 May 16, 2008 18:58

participant in question. The same pragmatic motivation is seen in relation to the patientive participant in the case of many antipassive expressions.

The passive and the antipassive present an interesting situation where the relevant nominals, despite the difference in their semantic roles, take on similar case forms. That the patient nominal in a passive clause occurs in the nominative case is an automatic consequence of its becoming a derived S or the sole argument of an intransitive clause due to suppression of the agent. In many languages, e.g. Quechua, the agent cannot be overtly expressed in their passive clauses. The passive agentive nominal, if allowed, takes on a case form used to mark an adjunct nominal of the basic clause type, i.e. an instrumental (Nepali *-bATa*, Russian *-m*, Kinyarwanda *n′-*) or locative case. The locative case may be a stationary locative (English *by*, North Russian *u* 'at'), an ablative (Nepali *-dwArA*, Ainu *orowa*, German *von*, French *de*), or an allative adpositional type (Japanese *ni*), which may be identical to the dative case in some languages (Japanese).

(8) Nepali (provided by Madhav Pokharel)
   a. *pulis-e ma-lAi jel-mA hAl-yo*
      police-ERG I-DAT jail-LOC PUT-3SG.PRF
      'The police put me in jail.'
   b. *pulis-dwArA/bATa ma jel-mA hAl-i-ẽ*
      police-ABL/INS I.NOM jail-LOC put-PASS-1SG.PRF
      'I was put in jail by the police.'

(9) Ainu (Shibatani 1990: 57)
   a. *Kamuy umma ray-ke.*
      bear horse die-CAUS
      'The bear killed a horse.'
   b. *Umma kamuy-orowa a-ray-ke.*
      horse bear-from PASS-die-CAUS
      'The horse was killed by a bear.'

The antipassive appears different from the passive in that the overt expression of a demoted patient seems to be the norm (except for the so-called absolute antipassive constructions of the form *Horses kick*, which express a generic or habitual property of a referent), while the passive agent does not typically surface even in those languages allowing such an agent.

(10) Chukchee (Kozinsky et al. 1988: 652)
   a. *əltəg=e keyŋ=ən penrə=nen*
      father=ERG bear=ABS attack=3SG:3SG/AOR
      'The father attacked the bear.'
   b. *əltəg= ən penrə=tko=gʔe keyŋ=etə*
      father-ABS attack=APASS=3SG.AOR bear=DAT
      'The father rushed at the bear.'

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 328 of 338 May 16, 2008 18:58

(11) Samoan (Mosel and Hovdhaugen 1992: 108)
a. *Sā ʻai e le teine le iʻa*
PAST eat ERG ART girl ART fish
'The girl ate the fish.'
b. *Sā ʻai le teine i le iʻa*
PAST eat ART girl LOC ART fish
'The girl ate some (of the) fish.'

As in the case change of the patient nominal in the passive clause, the change in case from the ergative to the absolutive in the agentive nominal of the antipassive clause is an automatic consequence of the change from its A status to the derived S status. As noted above, the demoted patient in an antipassive clause, just like the demoted agent in a passive clause, takes a case form that typically marks adjuncts in the basic clause type. Besides the dative and the locative form seen above, the demoted patient in the antipassive clause may be in an ablative (Central Yupik) or an instrumental form, as in Warrungu:

(12) Warrungu (Tsunoda 1988: 598)
a. *pama-ngku kamu pitya-n.*
man.ERG water.ABS drink-NFUT
'A man drank/drinks water.'
b. *pama kamu-ngku pitya-kali-n*
man.ABS water-INS drink-APASS-NFUT
'A man drank/drinks water.'

Notice that the Warrungu ergative case is identical with the instrumental case. Tsunoda (1988: 600–1) points out that instrumental marking of the antipassive patient is associated with those low in the animacy hierarchy, i.e. typically with an inanimate patient and occasionally with a human or other animate patient, but never with pronouns. This is also similar to the pattern of ergative marking across languages. Dative marking of the antipassive patient, on the other hand, occurs with all types of patient along the animacy hierarchy, according to Tsunoda (1988).

Antipassive clauses in ergative languages typically involve change in case marking and in verbal morphology, but the Samoan example (11b) above instantiates the pattern where a change in case marking alone signals the active/antipassive opposition.

21.3 Antipassive in accusative languages

Whether accusative languages also have an antipassive is as controversial as the question about the presence of a passive in ergative languages. As for the former, the

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 329 of 338 May 16, 2008 18:58

question boils down to whether the following kinds of contrast are to be accounted for in terms of the notion of voice:

(13) Japanese
a. *Taroo-wa maiban biiru-o nom-u.*
Taro-TOP every.night beer-ACC drink-PRES
'Taro drinks beer every night.'
b. *Taroo-wa maiban nom-u.*
Taro-TOP every.night drink-PRES
'Taro drinks every night.'

(14) English
a. *John shot the deer.*
b. *John shot at the deer.*

If the antipassive is defined functionally, as we do for other voice categories in this chapter, then the (b) sentences above would count as antipassive, for a patient in them either has low discourse relevance (13b) or it has not been (totally) affected (14b). At the formal level, the patient nominal is syntactically demoted and is in an oblique form, as in (14b), just like typical antipassive clauses. Unlike the 'true' antipassive, however, the relevant clauses do not change the case of the agent, nor do they involve verbal morphology indicating the derived clause status. The former property is due to the fact that in accusative languages the A nominal of a transitive clause and the (derived) S nominal are identically marked nominative. The question remains, however, whether the A nominal of a transitive clause changes its function to that of the derived S nominal under antipassivization here. Examination of the Japanese causative patterns indicates that other than the cases involving the verb *taberu* 'eat', the antipassive clause of the type seen in (13b) remains transitive with the A nominal retaining its A status. Compare the following:

(15) Japanese
a. *Kodomo-tati-ga/wa (takusan) tabe-ru.*
child-PL-NOM/TOP (a lot) eat-PRES
'Children eat (a lot).'
b. *kodomo-tati-o tabe-sase-ru*
child-PL-ACC eat-CAUS-PRES
'to make children eat'

(16) a. *Buka-tati-ga/wa (maiban) nom-u.*
subordinate-PL-NOM/TOP (every night) drink-PRES
'The subordinate (colleagues) drink (every night).'
b. *buka-tati-*o/ni nom-ase-ru*
subordinate-ACC/DAT drink-CAUS-PRES
'make the subordinate (colleagues) drink'

c. *buka-tati-*o/ni* *biiru-o* *nom-ase-ru*
subordinate-ACC/DAT beer-ACC drink-CAUS-PRES
'make the subordinate (colleagues) drink beer'

Under Japanese causativization, the subject (the causee nominal) of an intransitive clause takes the accusative *-o* (or the dative *-ni*), while that of a transitive clause exclusively takes the dative *-ni* marker (while the original patientive nominal retains the accusative marker – see 16c). As seen above, the antipassive form of *taberu* 'eat' (15b) follows the intransitive pattern, while that of *nomu* 'drink' follows the transitive pattern (16b).

Languages with a rich case inventory may also express antipassive semantics by means of case. Estonian, along with Finnish, is well known for its use of the partitive case in signalling an unaffected or partially affected patient, as in the following examples, where the affectedness status of the patient is marked in terms of the genitive/partitive case contrast. Russian and some other Slavic languages do the same with the accusative/genitive contrast.

(17) Estonian (provided by Kazuto Matsumura)
a. *Kunstnik joonistad pildi.*
artist paint picture.GEN
'The artist paints a picture.'
b. *Kunstnik joonistad pilti.*
artist paint picture.PART
'The artist is painting a picture.'

(18) Russian (provided by Elena Maslova and Vera Poddleskaya)
a. *On vypil piv-o.*
he drank beer-ACC
'He drank the beer.' (the whole available amount)
b. *On vypil piv-a.*
he drank beer.GEN
'He drank some beer.'

21.4 Passive in ergative languages

Turning now to the question of a passive in ergative languages, those languages that organize transitive constructions according to the absolutive–ergative case marking pattern may also have a passive without verbal morphology. Samoan form (19b) below illustrates what appears to be a typical way of expressing the functional equivalent of the passive in ergative languages:

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 331 of 338 May 16, 2008 18:58

(19) Samoan (Cook 1988: 30, 49)
a. *Na opo e le tama le teine.*
PAST hug ERG the boy the girl
'The boy hugged the girl.'
b. *Na opo le teine.*
PAST hug the girl
'Someone hugged the girl/ The girl was hugged.'
c. **Na opo e le tama.*
PAST hug ERG the boy
'The boy hugged.'

The pattern above is strikingly similar to the active/antipassive one in accusative languages (see Japanese examples in (13) above) in that in both cases the arguments bearing low discourse relevance are simply suppressed. Just as in the case of the agent of an antipassive clause in an accusative language, the patient of a passive clause in an ergative language does not change its case because of the identical absolutive marking for the P of a transitive clause and the (derived) S nominal. The question of whether or not the promotion of P to S actually takes place – i.e. whether a clause such as (19b) is intransitive (involving the status change of the patient nominal from P to S) or is still transitive (with the patient in the P function) – is again rarely raised in connection with the kind of passive under discussion. A literature survey indicates that there are both cases here too. The Mam forms in (20) below show a pattern similar to the Samoan examples above. Here too the agent in the passive clause in (b) is removed from the ergative A position (with concomitant removal of the ergative marking in the verb), and is optionally appended as an oblique headed by a relational noun.

(20) Mam (England 1988: 532, 537)
a. *o chi tzaj t-tyu-7n Xwan xiinaaq*
PAST 3PL.ABS DIR 3SG.ERG-grab-DIRS Juan man
'Juan grabbed the men.'
b. *ma ch-ok b'iyo-7n kab' xiinaq (t-u7n Cheep)*
REC 3PL.ABS-DIR hit-DIRS two man 3SG-RN/AG Jose
'The men were hit (by Jose).'

The Mam clause (20b) takes a directional (*ch-ok*) and is marked by the directional suffix -7*n*, which 'accompanies other transitive verbs with directionals, but [which] does not co-occur with intransitives or other passives with directionals', according to England (1988: 536). This shows that the clause in (20b) retains a transitive character; that is, the patient nominal has not converted to S.

Compared to the Mam case above, the S status of the patient in the Burushaski passive clause in (21b) below is clear from the agreement marker in the auxiliary.

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 332 of 338 May 16, 2008 18:58

(21) Burushaski (Morin and Tiffou 1988: 500)
a. *ne* *hír-e* *phaló* *bók-um* *bá-i*
the.MASC man-ERG seed.PL.ABS SOW.PRET-PTCP be-3SG.M.SBJ
'The man has planted the seeds.'
b. *phaló* *bók-um* *b-icá*
seed.PL.ABS SOW.PRET-PTCP be-3PL.SBJ
'The seeds have been planted.'

21.5 Case hierarchy and verbal voice morphology

Before zeroing in on the correlation between case marking patterns and verbal voice marking, let us briefly touch on the issue of the primary argument in a nominative–accusative and an absolutive–ergative clause. Discussions (e.g. Anderson 1976) of grammatical relations that give substantially more weight to syntactic behavioural properties than to morphological coding patterns recognize the equivalence in the grammatical role between the nominative subject and the ergative agentive nominal, and assign the grammatical relation of subject to both types of nominals. Such a move is questionable in view of the fact that the ergative agentive nominal is not entirely like the nominative subject of accusative languages. For one thing, ergative agentive nominals are freely omissible, yielding those passive clauses just examined above; but this is not usually the case with the nominative subjects of accusative languages. While some phenomena (e.g. the control of reflexives and certain types of gap in coordinate and subordinate clauses) point to the similarity between the ergative agentive nominal and the nominative subject, others (especially the markedness in case form and omissibility) point to the similarity between the ergative agentive nominal and the accusative object. The controversy over the question of the grammatical relations of the absolutive and ergative nominals notwithstanding, the equivalence of case function between the nominative and the absolutive case is clear. Both are typically unmarked and identify the indispensable nominal. These two criteria establish the following case hierarchies for the nominative–accusative and the absolutive–ergative case marking systems (cf. Woodbury 1977: 330, Blake 1994: 90):

(22) Case hierarchies
NOM > ACC > DAT… (Accusative languages)
ABS > ERG > DAT… (Ergative languages)

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 333 of 338 May 16, 2008 18:58

These case hierarchies help us gain some understanding of the correlation between case systems and verbal voice marking. The observations above indicate that when a voice alternation affects a secondary case form, as in the antipassive in accusative languages and the passive in ergative languages, the relevant argument in a secondary case form is omitted from the clause or demoted to an oblique status – see (13)–(14) and (19)–(21). In this kind of voice operation, the primary case form is not disturbed. Often the resulting clause remains transitive, where the nominative A or absolutive P nominal remains intact. Even if the nominative A or the absolutive P nominal turns to S under intransitivization, its case marking does not change. It is under these circumstances, where the unmarked primary case form remains constant, that verbal voice morphology is typically absent. When the voice alternation affects an unmarked primary case form, as in the passive in accusative languages and the antipassive in ergative languages, verbal voice marking typically occurs. The constraint that a clause must in principle contain an unmarked case form (see Shibatani 1977 and Tsunoda 1981b) typically necessitates a promotion of a secondary argument to the derived-S role associated with an unmarked case when the basic unmarked case form is removed under voice alternation. The verbal morphology associated with this type of voice alternation signals a more radical structural reorganization involved in it (see Malchukov 2006a for a similar analysis).

The generalization above certainly does not preclude the following two possibilities: passives and antipassives without promotion of a secondary argument and the possibility of verbal marking in those voice alternations affecting a secondary argument. The former cases are rarer than promotional passives and antipassives, while antipassive verbal marking in accusative languages (as in Russian below) and passive verbal marking in ergative languages (as in the Mam example below) seem fairly widespread.

Nonpromotional passive (cf. the promotional passive in (8b) above)

(23) Nepali (provided by Madav Pokharel)
ma-lAi jel-mA hAl-i-yo
I-DAT jail-LOC put-PASS-3SG.PRF
'(They) put me in jail.'

(24) Warlpiri (Dixon 1980: 449)
a. *nyuntulu-rlu Ø-npa-ju pantu-rnu ngaju*
you-ERG Ø-2SG.A-1SG.P spear-PAST I.ABS
'You speared me.'

Nonpromotional antipassive

b. *nyuntulu-rlu Ø-npa-ju-rla pantu-rnu ngaju-ku*
you-ERG Ø-2SG.A-1SG-DAT spear-PAST I-DAT
'You speared at me; you tried to spear me.'

(25) Russian (provided by Elena Maslova and Vera Poddleskaya)
a. *Babuška rugajet vnuka.*
granny scolds grandson.ACC
'Granny is scolding the grandson.'
b. *Babuška rugajet-sja.*
granny.NOM scold-APASS
'Granny is scolding.'

(26) Mam (England 1988: 535)
ma Ø-juusa-njtz chib'aj t-u7n Mal
REC 3SG.ABS-burn-PASS food 3SG-RN/AG Maria
'The food was burned by Maria.' (by accident)

21.6 Applicative

Due to space limitations, only one type of valency-increasing phenomenon, namely the applicative voice, and some related issues can be discussed here. Applicatives have the syntactic function of introducing an additional patient-like entity into the argument structure, which is converse of the passive and antipassive syntactic function, which removes an argument from the core-argument structure. Conceptually, the applicative voice signals the presence of an extra affected entity (in addition to the normal patient, in the case of a transitive situation). Thus, Balinese locative applicative form (27b) below implies that the entire garden has been planted with bananas, and Walihío benefactive (28b) implies that the speaker came into possession of the cooked beans, while the non-applicative counterparts have no such implications. The pragmatic effect of the applicative voice is seen in the Amharic examples in (29), where the applied status of an instrumental argument signals its higher discourse relevance.

(27) Balinese (Arka 2003)
a. *Tiang mulan biyu di tegalan tiang-e*
1SG plant banana in garden 1SG-POSS
'I planted bananas in my garden.'
b. *Tiang mulan-in tegalan tiang-e biyu*
1SG plant-APPL garden 1SG-POSS banana
'I planted my garden with bananas.'

(28) River Walihío (Felix 2005)
a. *hustína pasu-ré muní no'ó ičió*
Agustina cook-PFV beans 1SG.NONNOM BEN
'Agustina cooked beans for me.'

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 335 of 338 May 16, 2008 18:58

b. *hustína no'ó pasú-ke-re muní*
Agustina 1SG.NONNOM cook-BEN-PFV beans
'Agustina cooked me beans.'

(29) Amharic (provided by Mengistu Amberber)
a. *aster siga bə-tillik' billa k'orrət'ə-čč*
Aster meat with-big knife cut.PF-3F
'Aster cut some meat with **a big knife**.'
b. *aster tillik'-u-n billa siga k'orrət'ə-čč -ibba-ət*
Aster big-DEF-ACC knife meat cut.PF-3F-APPL-3MO
'Aster cut some meat with **the big knife**.'

(30) Diyari (Austin 1981: 159)
a. *n̪awu kanku piṭa-n̪i ṯika-yi*
3SG.NF.S boy stick-LOC return-PRES
'The boy is going back with a stick.'
b. *n̪ulu kanku-yali piṭa ṯika-lka-yi*
3SG.NF.A boy-ERG stick-ABS return-APPL-PRES
'The boy is taking a stick back.'

As far as the case marking pattern goes, the promoted applied nominal takes on the case form of the patient in the language. Thus, if the language uses a bare form for its patient, then the applied nominal is also in a bare form, as in Balinese; if the language makes a two way nominative/non-nominative distinction, the applied nominal will be in the non-nominative case, as in River Walihío; and if the language has a distinct accusative case, the applied nominal triggers the accusative marking, as in Amharic. In ergative languages, the applied nominal again will be treated like a patient and is realized in the absolutive case form, while the agent of an applicative clause assumes the ergative case, as in the Diyari example (30b).

In languages that make a distinction between the accusative and the dative case, the beneficiary nominal of a benefactive applicative generally takes the dative case. Korean and Japanese, however, differ in this respect. Korean allows accusative marking on the applied beneficiary, while Japanese does not. This is apparently connected to the fact that while Korean allows double accusative constructions with the basic ditransitive verb 'give' (*cwuta*), Japanese does not.

(31) Japanese
a. *Hahaoya-ga kodomo-ni/*-o hon-o yat-ta.*
Mother-NOM child-DAT/-ACC book-ACC give-PAST
'Mother gave a book to the child.'
b. *Hahaoya-ga kodomo-ni/*-o hon-o kat-te yat-ta.*
mother-NOM child-DAT/-ACC book-ACC buy-CON GIVE-PAST
'Mother bought the child a book.'

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 336 of 338 May 16, 2008 18:58

(32) Korean
a. *emeni-ka ai-eykey/-lul chaek-ul cwu-ess-ta.*
mother-NOM child-DAT/-ACC book-ACC give-PAST-IND
'Mother gave a book to the child.'
b. *emeni-ka ai-eykey/-lul chaek-ul sa-cwu-ess-ta.*
mother-NOM child-DAT/-ACC book-ACC buy-GIVE-PAST-IND
'Mother bought the child a book.'

As opposed to the earlier examples in which both case and verbal morphology signal the applicative voice, the Japanese and Korean forms above show a situation in which the applicativization is solely marked by verbal morphology. These constructions are, in fact, serial verb constructions in OV languages, where a verb (e.g. GIVE) introduces as its object a nominal (e.g. a beneficiary nominal) not subcategorized for by the main verb. Serial verbs of this type are widely seen among those languages with an impoverished case system, where a serialized verb may grammaticalize and become an adposition. This development is seen more clearly in a VO language such as Chinese (where, as in many verb serializing languages, the passive and causative constructions also take the form of serial verbs):

(33) Chinese (Shibatani, Zhāng and Lú 1994: 460, 464)
a. *Wǒ gěi háizi yì běn shū.*
I give child one CL book
'I gave the child a book.'
b. *Wǒ zuò fàn gěi háizi.*
I cook rice GIVE child
'I cooked the child rice.'

While *gěi* in (33b) above is preposition-like, it is important to recognize its role as a marker of the benefactive applicative as the dative case form in the German benefactive construction in (34a) rather than as the prepositional counterpart in (34b) below:

(34) German
a. *Ich kochte dem Kind Reis.*
I cooked the.DAT child rice
'I cooked the child rice.'
b. *Ich kochte Reis für das Kind.*
I cooked rice for the.NOM child
'I cooked rice for the child.'

In (33b) and (34a), the beneficiary is clearly more involved in the event as an intended recipient of the cooked rice as in the typical transitive-based benefactive applicative construction in other languages.

So-called external possession constructions typically follow the structural templates of benefactive applicatives since they too express a voice where the action extends beyond the patient and affects an additional entity, typically construed as the owner of the patientive reference, as in the following examples:

(35) German
Sie wäscht dem Kind die Haare. (cf. 34a)
she washes the.DAT child the hair
'She washes the child's hair.'

(36) Korean
emeni-ka ai-lul son-ul ttayli-ess-ta. (cf. 32b)
mother-NOM child-ACC hand-ACC hit-PAST-IND
'Mother hit the child on the hand.'

Among Formosan and Philippine-type Western Austronesian languages that preserve the proto-Austronesian focus system to varying degrees, applicativization and focusing (or voice selection) combine such that applied nominals automatically become the topic (or subject) of the clause. Hence, the applied nominals in such oblique-centred focus constructions as the locative, the benefactive, and the instrumental focus construction are the topics/subjects of the respective clauses in the nominative case, as the following examples in (37c–d) from the near-extinct Formosan language Pazih:

(37) Pazih (Li and Tsuchida 2001: 30)
a. *me-xe'et nuang ki yaku.* (Actor-focus/voice)
AF-tie cow NOM I
'I tied a cow.'
b. *xe'ed-en naki lia ki muang.* (Patient-focus/voice)
tie-PF I.GEN ASP NOM cow
'The cow has been tied by me.'
c. *saa-xe'et nuang ki saris.* (Instrumental-focus/voice)
IF-tie cow NOM string
'The string was used to tie a cow.'
d. *x<in>e'ed-an nuang ki imini a kahuy.* (Locative-focus/voice)
PRF-tie-LF cow NOM this LIG tree
'The tree was the place where the cow was tied.'

Notice that while the patient-focus form has been translated as passive in (34b), the agentive nominal is characteristically expressed in such a clause, unlike in the typical passive clause. Both instrumental and locative focus forms are monoclausal, unlike the accompanying English translations. All four forms above are voice variants of the same verb form.

The coupling of applicativization and topicalization/subjectivization in these Austronesian languages is a natural consequence of the functional underpinnings

978–0–19–920647–6 22-Spencer-ch21 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 338 of 338 May 16, 2008 18:58

of voice phenomena in these languages. Formosan and Philippine-type Western Austronesian languages generally select the patient-focus form (e.g. 37b above) when the patient has a high degree of conceptual or discourse relevance (see McFarland 1978, Shibatani 1988, Nolasco 2005). Applied nominals express an entity of high relevance by virtue of its affectedness status and/or its discourse relevance. Since an applied nominal is treated like a patient nominal of the language, it will be automatically topicalized in these Austronesian languages.

CHAPTER 22

DIFFERENTIAL CASE MARKING AND ACTANCY VARIATIONS

ANDREJ MALCHUKOV
PETER DE SWART

22.1 Transitivity and Case Alternations

In the early 1980s Hopper and Thompson (1980) and Tsunoda (1981b) independently of each other argued in favour of a prototype approach to transitivity in which transitivity is treated as a gradable and multi-factorial notion (see also Moravcsik 1978 for a pioneering study of object alternations). A high or prototypical transitive clause correlates with the presence of a number of parameters, or in Tsunoda's terms satisfaction of the Effectiveness Condition (EF-CON). Such a prototypical transitive clause is realized with a transitive case frame: NOM–ACC or ERG–ABS with one marked (ACC/ERG) and one unmarked (NOM/ABS) case (depending on the alignment type). In particular, Hopper and Thompson have identified the parameters shown in Table 22.1 contributing to transitivity of a construction.

Table 22.1. Transitivity Parameters (Hopper and Thompson 1980)

Parameter	High	Low
A. Participants	2	1
B. Kinesis	action	non-action
C. Aspect	telic	atelic
D. Punctuality	punctual	non-punctual
E. Volitionality	volitional	non-volitional
F. Affirmation	affirmative	negative
G. Mode	realis	irrealis
H. Agency	A high in potency	A low in potency
I. Affectedness of	O totally affected	O not (totally) affected
J. O individuation	O highly individuated	O non-individuated

Both papers demonstrate that absence of certain semantic parameters or failure of the EF-CON can result in a transitivity alternation involving such morphosyntactic mechanisms as agreement, incorporation, and diathetic shifts (e.g. antipassivization). The effect of these parameters on case morphology can also be readily demonstrated, where changes in case morphology are referred to as *case alternations*. Restricting ourselves to the parameters proposed by Hopper and Thompson (1980; hereafter H&T)[1], we find, for instance, that in Georgian and Hindi tense/aspect (C) condition a case split (e.g. Harris 1981; Mohanan 1994). Volitionality (E) and agency (H) often influence case marking, a good example being the ERG–ABS alternation on the S-argument in fluid-S languages such as Bats (Holisky 1987; see Dixon 1994 for additional discussion). In Russian and other Slavic languages ACC on the O changes to GEN under negation (F) (see de Hoop and Zwarts, this volume, for discussion). Partitive (as in (8) below) instead of ACC/GEN is used on Estonian O-arguments when the O is only partially affected (I). Finally, the parameter O-individuation (J) is a cover term for features such as animacy and definiteness. The discussion of *differential object marking* below shows how they play a role in case alternations on objects. (See Kittilä, Chapter 23, for further discussion of the relation between transitivity parameters and transitivity alternations.)

One limitation of H&T's approach is that it does not constrain co-variation between different types of transitivity parameters and transitivity alternations (subject alternations, object-alternations, passive and antipassive derivation, incorporation, and the like). In the subsequent literature this co-variation has been shown to be subject to additional constraints (see section 22.6 below).[2]

[1] Parameters are referred to by means of their letter in the Hopper and Thompson article.

[2] Another aspect of H&T's approach which is open for further discussion concerns their 'Transitivity Hypothesis', which predicts systematic co-variation between individual transitivity

978–0–19–920647–6 23-Spencer-ch22 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 341 of 355 May 16, 2008 18:58

22.2 Transitivity alternations and transitivity splits

While H&T restrict themselves to transitivity alternations, i.e. the same verb occurs with different case frames depending on the transitivity of the clause, Tsunoda (1981b, 1985) argues that the same principles (his EF-CON) can be used to describe verb splits, that is, different lexical classes of verbs select for different case frames. This is indeed a sensible approach given that what surfaces as a split in one language takes the form of an alternation in another. Consider the class of perception verbs: while in many languages active vs. inactive perception verbs are represented by different lexemes subcategorizing for different case frames (cf. 'see' vs. 'look at' in English), in other languages this verb is involved in a transitivity alternation; thus in Even (Tungusic), the same verb *it*= means 'see' when combined with an accusative object and 'look' when combined with an allative object. On Tsunoda's approach *see*-verbs unlike *look*-verbs imply impingement of action on O, which would account for the preference for transitive encoding (impingement is a parameter related to H&T's affectedness).

We can thus distinguish two main classes of case alternations: 1) split alternations, displaying an alternation of lexical case associated with different verbal lexemes; 2) fluid alternations, when the same verb takes alternative case frames depending on transitivity parameters (cf. de Hoop and Malchukov 2006; the terms split vs. fluid are adopted from the Dixonian distinction between split-S and fluid-S; Dixon 1994). Furthermore, we can make a distinction between *differential subject marking*, in case the alternation occurs on the subject argument, and *differential object marking*, in case of an alternation on the object argument.

On the formal side a distinction can be made between (i) *asymmetric case alternations* in which the alternation occurs between zero expression of case, i.e. unmarked case, and overt case marking, e.g. an ABS–DAT alternation; and (ii) *symmetric case alternations* in which two overt cases alternate, e.g. ERG and DAT.[3]

features. In its strong form this hypothesis is controversial. As noted by Tsunoda (1985; cf. Lazard 1998), some parameters indeed co-vary (e.g. agentivity and animacy of the A participant), while some others do not (e.g. agentivity and affectedness). In a recent paper, Lazard (2003) questions the validity of the transitivity hypothesis in general, denying any systematic co-variation between parameters. A more moderate view is to assume a weaker form of the Transitivity Hypothesis that predicts that only parameters that are semantically related will show a systematic co-variation (Malchukov 2006a).

[3] The equation of unmarked case with zero exponence is a simplification inasmuch as it presupposes an iconic relation between form and function to the effect that functionally unmarked forms will be formally unmarked as well. This iconic relation is indeed a majority case even though some languages go against this tendency (e.g. some marked nominative languages, discussed by König, Chapter 35).

978–0–19–920647–6 23-Spencer-ch22 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 342 of 355 May 16, 2008 18:58

22.3 Split Case Alternations and Verb Type Hierarchies

Tsunoda (1981b, 1985) proposed the following verb type hierarchy which predicts the distribution of (in)transitive case frames in individual languages.

(1) Effective action ≫ Perception ≫ Pursuit ≫ Knowledge ≫ Feeling ≫ Relation

The hierarchy extends from the most transitive verbs (in terms of satisfaction of EF-CON) on the left to less transitive verbs on the right predicting that if a verb type lower on the hierarchy occurs with a transitive case frame so do all verb types higher on the hierarchy. Tsunoda provided cross-linguistic data in support of the hierarchy (see table 3 in Tsunoda 1981b). In particular, he shows that while some languages (like Eskimo) extend the transitive frame all the way down the hierarchy, other languages show earlier cut-off points. Thus in Djaru (Australian) effective action and perception predicates pattern transitively, while for lower types the transitive ERG–ABS frame alternates with an intransitive one, and in Avar (Daghestanian) even perception verbs pattern intransitively (taking an OBL–ABS frame). The following examples illustrate the early break-down of the transitivity pattern in Avar which in this respect is representative of Daghestanian languages (see Daniel and Ganenkov, Chapter 46).

Avar (Blake 2001: 121 from Ebeling 1966)

(2) *Inssucc-a* *j-as* *je-cc-ula*
(M)father-ERG F-child F-praise-PRES
'The father praises the girl'

(3) *Inssu-du* *j-as* *j-ix-ula*
(M)father-LOC F-child F-see-PRES
'The father sees the girl'

(4) *Inssu-je* *j-as* *j-óλ'- ula*
(M)father-DAT F-child F-love-PRES
'The father loves the girl'

While being credited as a major contribution to the typology of verb classes, Tsunoda's hierarchy has also been confronted with a number of problems (cf. Drossard 1991; Lehmann 1991; Lazard 1998). The hierarchy seems to conflate several semantic dimensions (Lehmann 1991: 234) and a strict ordering of intermediate verb types seems to be questionable (Lazard 1998) as some verb types are ranked one way in some languages and the other way in a different language. This holds, in particular, for the ordering of mental verbs (perception and emotion) with respect to pursuit verbs.

978–0–19–920647–6 23-Spencer-ch22 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 343 of 355 May 16, 2008 18:58

Malchukov (2005) tries to overcome these problems by decomposing Tsunoda's hierarchy along the two dimensions of decreased patienthood of O and decreased agenthood of A and recasting it in terms of a two-dimensional hierarchy which can be interpreted as a semantic map (see Haspelmath 2003 for a discussion of the semantic map approach as well as Malchukov and Narrog, Chapter 34).

(5) ≫ contact ≫ pursuit ≫ (motion) Effective ≫
≫ perception/cognition ≫ emotion ≫ (sensation) action

This two-dimensional hierarchy acknowledges the fact that different semantic dimensions play a role in the association of case frames with verb types. Moreover, this reformulation has the clear advantage that it can handle the instability in the ranking of the intermediate verb types such as the pursuit and mental verbs. As the two verb types are elements on different hierarchies they have no relative ranking with respect to each other. It can also better capture variations between languages related to differences in the extension of the transitive pattern along the two dimensions of the hierarchy. For example, Japanese, by treating pursuit predicates as transitive (e.g. *matu* 'wait for'), allows a broader extension of the transitive pattern along the first sub-hierarchy than English does. On the other hand, English is more liberal than Japanese in extending the transitive pattern along the second sub-hierarchy as it assimilates mental verbs to the transitive pattern, while Japanese switches to a DAT–NOM or NOM–NOM pattern (see Jacobsen 1992; Shibatani 2001):

Japanese (Jacobsen 1992: 30, 31)

(6) *(Watashi ni) kokuban ga mieta*
(I DAT) blackboard NOM see/visible-PAST
'I saw the blackboard'

Another limitation of Tsunoda's analysis is that while it predicts that some verb types lower on the hierarchy might deviate from the transitive case frame, it does not predict which case frame occurs with which verb type even though there exist clear cross-linguistic patterns. For example, in German both pursuit verbs (cf. *warten auf* 'wait for') and mental verbs (e.g. *gefallen* 'like') can be intransitive but only the latter can take the inverse DAT–NOM pattern. And the same is true for Ingush, an ergative language in which pursuit verbs (e.g. *hež* 'wait') take an ABS–OBL pattern, while mental verbs (e.g. *viez* 'like') take the DAT–ABS frame (Nichols 1994b: 118–19). Thus in both cases, mental verbs allow for an 'inverse' (DAT–NOM or DAT–ABS) pattern, while pursuit verbs do not. Malchukov (2005) shows how these patterns can be derived through a competition between a few functional constraints.

978–0–19–920647–6 23-Spencer-ch22 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 344 of 355 May 16, 2008 18:58

22.4 Constraints on Case Frames: converging approaches

Multi-factorial approaches to transitivity as advocated by Tsunoda and H&T converge with Dowty's (1991) work on proto-roles (see e.g. Ackerman and Moore 2001; Levin and Rappaport Hovav 2005 for discussion). Indeed many of Dowty's proto-role properties contributing to Proto-Agent and Proto-Patient roles find straightforward analogues among H&T's transitivity parameters, or factors contributing to Tsunoda's EF-CON (see Primus, this volume, for discussion of Dowty's approach). Yet it should be kept in mind that originally Dowty's approach has been designed to account not for transitivity alternations as such, but rather as a mechanism governing argument selection. The main insight is that the argument which displays more Proto-Agent properties will be encoded as the subject. This approach provides a natural account for a cross-linguistic consistency in the encoding of semantic transitives (where agent and patient map unproblematically to subject and object, respectively). It also accounts for cross-linguistic variation in the domain of experiencer predicates which show a split between subject-experiencer (*I fear it*) and object experiencer constructions (*It frightens me*) (see Ackerman and Moore 2001: 62–71). The two arguments of experiencer predicates have an equal number of Proto-Agent properties, the experiencer being sentient and the stimulus being a causer. As a result either argument is entitled to be encoded as the subject which is reflected in cross-linguistic variation.

This line of research has been extended to account for transitivity alternations as well. Thus, Primus (1999) accounts both for a switch in subject encoding from NOM (or ERG) to DAT and for subject encoding in languages with split intransitivity in terms of Proto-Agent and Proto-Patient properties (see also Primus, Chapter 17). In a similar fashion, Ackerman and Moore (2001) suggest complementing Dowty's 'Syntagmatic Selection Principle' with a 'Paradigmatic Selection Principle'. The latter accounts for the observation that when a core argument alternates with an oblique, the latter shows a decrease in proto-role properties. For example, in Estonian, direct objects can alternate between accusative and partitive where the latter signals absence of the Proto-Patient property *bounding entity* introduced by Ackerman and Moore (2001, chapter 5).

Estonian (Ackerman and Moore 2001: 109)

(7) *Madis joob teed*
Madis drink.3SG.PRES tea.PART
'Madis is drinking tea'

(8) *Madis joob oma tee ära*
Madis drink.3SG.PRES own tea.GEN/ACC PREV
'Madis will drink up his tea'

978–0–19–920647–6 23-Spencer-ch22 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 345 of 355 May 16, 2008 18:58

In the functional-typological tradition such alternations are discussed in terms of affectedness. Thus, the Dowtian line of research is extended to the domain of transitivity alternations, as pioneered by Moravcsik, Hopper and Thompson, and Tsunoda. Independently, cross-linguistic variation in case marking has been investigated from an optimality-theoretic perspective (see de Hoop, Chapter 6). In the OT tradition much of the variation in the encoding of different verb classes (e.g. experiencer verbs), is conceived in terms of competition of semantic case assignment with structural case assignment and/or markednes/economy constraints. Later developments have shown further convergence of different research traditions in the study of transitivity alternations. (See de Hoop, Chapter 6, for further discussion.)

22.5 Fluid Case Alternations: Differential Object Marking

Coming back to fluid case alternation, we consider in this section the best studied phenomenon in this domain, known as differential object marking. In a language with differential object marking (DOM) one set of direct objects is case marked in one way and another set in a different way depending on features of the object.[4] The term DOM is most frequently used for languages which show an asymmetric case alternation (cf. Bossong 1985a, 1991; Aissen 2003). The phenomenon does not only surface in case marking but can also result in variation in agreement or word order. According to Bossong (1985a) over 300 languages in the world show a DOM system of some sort.

Many of the features influencing DOM can be related to Hopper and Thompson's notion of O-individuation. Recurrent dimensions are the animacy, definiteness, or specificity of the direct object. An example of animacy-based DOM is found in Malayalam (Asher and Kumari 1997), where animate objects may be marked. Consider the following examples:

Malayalam (Asher and Kumari 1997:203)

(9) *Avan oru paΣuvine va‖i*
he a cow.ACC buy.PAST
'He bought a cow'

[4] The notion of case marking is used here as a broad concept also referring to adpositional marking of objects.

(10) *Avan pustakam vaayiccu*
He book read-PAST
'He read the book'

A similar DOM pattern is found in Guaraní (Bossong 1985b). Definiteness plays a role in Persian (Lazard 1992) and Amharic (Amberber 2005 and Chapter 51) in which only definite objects are case-marked. For Turkish, it has been reported that specificity plays an important role in the marking of objects (see Johanson 2006 *et passim*; Von Heusinger and Kornfilt 2005 for discussion). Combinations of features, referred to as two-dimensional DOM in Aissen (2003), are also attested as in the DOM systems of Hindi (e.g. Mohanan 1994) and many Romance languages (e.g. Spanish, Rumanian, Sardinian; see Bossong 1991 for references) in which both animacy and definiteness/specificity play a role. An exceptional case is found in Palauan in which the marking of objects is determined by animacy, specificity, and number interacting with the aspectual system (Woolford 1995).

Finally, there are languages in which NP type seems to play a role as in the Australian language Pitjantjatjara in which only pronouns and proper nouns are case-marked. Often NP type is conflated with the notion of definiteness (cf. Aissen 2003). Another notion closely related to definiteness and specificity is that of topicality (or even broader information structure) which is argued to influence DOM in some languages (see Dalrymple and Nikolaeva 2006).

As noted by Bossong (1985a) DOM most typically occurs in nominative–accusative languages. Examples of ergative languages with DOM are nevertheless attested as the following example from Warlpiri demonstrates:

Warlpiri (Hale 1973)

(11) *Njuntulu-lu npa-tju pantu-nu ŋatju*
2SG-ERG 2SG-1SG spear-PAST 1SG.ABS
'You speared me'

(12) *Njuntulu-lu npa-tju-la pantu-nu ŋatju-ku*
2SG-ERG 2SG-1SG-la spear-PAST 1SG-DAT
'You tried to spear me'

Here, as in a more literal English translation of (12) 'You speared at me', a less affected O goes into an oblique, while ERG marking is retained. In most ergative languages, however, the ERG–OBL pattern is disallowed, and demotion of an O in a would-be DOM pattern triggers a shift to an antipassive construction, where A surfaces as an absolutive S.

Warrungu (Tsunoda 1985)

(13) *Pama-ngku yuri nyaka-n*
man-ERG kangaroo-ABS see-NF
'A man saw (found) a kangaroo'

978–0–19–920647–6 23-Spencer-ch22 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 347 of 355 May 16, 2008 18:58

(14) *Pama yuri-wa nyaka-kali-n*
man.ABS kangaroo-DAT see-APASS-NF
'A man was (is) looking for a kangaroo'

Notably, as observed by Evans (1995a: 344), those (relatively few) Australian languages which like Kayardild have accusative alignment display DOM more regularly, as compared to languages with a (split) ergative alignment. We shall return to the explanation of this asymmetry between accusative and ergative languages in the domain of differential object marking in section 22.6 below.

Symmetric case alternations on objects are also attested although they are normally not discussed under the heading of DOM. A good example of such a system is the alternation from ACC/GEN to PART in Finnish under the influence of aspect (completedness) and affectedness of the object (cf. Kiparsky 1998 for a good discussion). Estonian shows a similar pattern (cf. Ackerman and Moore 2001, Tamm 2005 among others; see examples (7) and (8) above). The Russian genitive of negation falls in the same category (see de Hoop and Zwarts, Chapter 11, for discussion of this pattern). Remarkably, it seems that whereas the asymmetric DOM systems are typically dependent on features of the direct object (parameter of O-individuation in terms of H&T), the symmetric DOM systems often pertain to a broader range of parameters (factivity–aspect–affectedness–individuation).

22.6 Explanations for Case Alternations: Functions of Case Marking and Economy

In functional-typological literature two main functions of case marking have been identified: the indexing and the distinguishing function (Comrie 1989; Mallinson and Blake 1981; Kibrik 1985). Although initially seen as alternative and even competing motivations, both are now seen as necessary ingredients in order to account for the cross-linguistic variation in case marking (Song 2001).

Under the distinguishing function, case is used to distinguish between the core arguments subject and object. This function can be illustrated with the following example from Awtuw (Feldman 1986), where the object is obligatorily marked with accusative case if the object is equally high or higher than the subject in the animacy hierarchy:

Awtuw (Feldman 1986: 106)

(15) *tey tale-re yaw d-æl-i*
3F.SG woman-ACC pig FAC-bite-PAST
'The pig bit the woman'

978–0–19–920647–6 23-Spencer-ch22 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 348 of 355 May 16, 2008 18:58

(16) *tey tale yaw d-æl-i*
3F.SG woman pig FAC-bite-PAST
'The woman bit the pig'

As the higher element in animacy is normally interpreted as the subject, the use of case marking prevents an interpretation in which the woman is regarded as the subject in (15). The object marking here is clearly due to the relation between the subject and object and can therefore be referred to as *global* distinguishability. De Swart (2003, 2006) uses this notion of distinguishability to account for DOM patterns similar to that of Awtuw.

Many analyses of asymmetric DOM, however, relate the phenomenon to the distinguishing function of case marking in terms of *markedness reversal* (cf. Comrie 1989; Bossong 1991; Aissen 2003; see also de Hoop, Chapter 6). These approaches argue that those objects which look too much like prototypical subjects are marked in order to distinguish them from the subject. It should be noted that distinguishability is used here in a *local* sense as the object is not made distinct based on a comparison with the actual subject in the sentence but with a notion of prototypical subject.[5] The distinction between the two types of distinguishability is not always clear-cut. Whereas in Hindi DOM is best explained in terms of local distinguishability (cf. Aissen 2003), the genetically related Kashmiri is better treated in global terms. In Kashmiri P takes an object (ACC/DAT) case if A is lower than P on the Animacy/Person Hierarchy (Wali and Koul 1997: 155).

Kashmiri (Wali and Koul 1997: 87–8)

(17) *Su chu me parɪna:va:n*
he.NOM is I.OBJ teaching
'He is teaching me'

(18) *Nanan roT-u-s bɪ*
Nana.ERG caught-M.SG-1SG I.M.SG.NOM
'Nana caught me'

Similar 'global' constraints are at work in Yukaghir, where ACC marking of P is lacking if A is first or second person (Maslova, Chapter 55).

The two notions can also co-exist within one language: in Spanish and Malayalam, for instance, DOM is mostly local, but sometimes global (cf. de Swart 2003, 2006; Morimoto and de Swart 2006). Thus in Malayalam inanimate Os are generally unmarked, but they can take ACC marking in case the subject is also inanimate (see de Swart 2003 for discussion); compare (19) and (9)–(10) above.

[5] Cf. Silverstein's (1976) distinction of 'global' and 'local' case marking: the former differs from the latter in that case assignment is sensitive to properties of both arguments, rather than to properties of the host NP.

Malayalam (Asher and Kumari 1997: 204)

(19) *Kappal tiramaalaka⟩e bheediccu*
ship wave-PL.ACC split-PAST
'The ship broke through the waves'

The relation between global and local distinguishability can also be conceived in diachronic terms: what starts as a pattern of global distinguishability, where the use of a marker is optional (dependent on context, which can always make its use dispensable), is eventually conventionalized as a pattern of local distinguishability (where say animate objects are always marked irrespective of context).[6] This is consistent with what we know about the extension of DOM in individual languages (see e.g. Aissen 2003 for a diachronic discussion of DOM in Spanish; cf. also Morimoto and de Swart 2006). A related development occurred in Persian where an animacy-based DOM (attested in early Judaeo-Persian texts) developed into a definiteness-based pattern (see Stilo, Chapter 48). This conclusion is also compatible with analyses which suggest that 'pragmatic' DOM can eventually evolve into 'semantic' DOM (Zeevat and Jäger 2002; Jäger 2004), as well as with Baerman's conclusion that alternations which started as functionally motivated may eventually 'desyntacticize', i.e. become a morphologized relic of what was once an active syntactic rule (Baerman, Chapter 14).

Together with a notion of economy, which ensures that in every non-elliptical transitive sentence we find an NP in the unmarked case (NOM/ABS), the so-called primary argument (cf. Palmer 1994[7]), the distinguishing function of case can account for the well-known tendency both in accusative and ergative languages to leave a single argument (S) of an intransitive verb and one of the core arguments (A or O) of the transitive verb unmarked (cf. e.g. Daniel and Ganenkov, Chapter 46, on 'the nominative requirement' in the ergative Daghestanian languages).

An economy constraint can be found in one formulation or the other in many approaches of different theoretical persuasions (cf. Shibatani 1977; Dixon 1979; Tsunoda 1981b; Bobaljik 1993b; Primus 1999; Nakamura 1999; Wunderlich and Lakämper 2001; de Hoop and Narasimhan 2005). A special formulation is the Primary Argument Immunity Principle (Malchukov 2006a) which states that manipulation of the case marking of the primary argument should be avoided. This type of constraint explains why case alternations affecting the case of the unmarked NP are cross-linguistically disfavoured and normally result in a voice alternation, cf. (13) and (14) above. Furthermore, it can explain the observation (Bossong 1985a; Drossard 1991) that DOM is predominantly found in accusative languages (where A is the primary argument), whereas DSM is most often found in ergative languages (where O is the primary argument).

[6] For an insightful functional account of this development see Durie (1995).

[7] Cf. the notion of 'predication subject' in Lazard (1998). A predication subject displays the following properties: obligatoriness, zero case marking, control of verbs agreement (Lazard 1998: 110).

Some other case phenomena, however, cannot be treated under the distinguishing view and should be left to the indexing view in which case is used to encode semantic roles. Generally distinguishability alone cannot account for oblique cases. In some languages ('role-dominated languages', in terms of Foley and Van Valin 1984) marking of core arguments can also be better explained in semantic terms. For example, in Manipuri (Bhat and Ningomba 1997) only volitional agents take ergative case while only affected patients take accusative case. More generally, distinguishability cannot account for patterns of symmetric case alternations, such as the ACC–PART alternation in Finnish and Estonian: indeed both cases satisfy distinguishability here, so their variation should have a semantic basis.

On the other hand, it is more difficult to handle cases of global case marking from an indexing perspective. This is so because indexing is basically a local strategy: iconicity would predict that other things being equal the role-related properties would be marked on the NP to which the properties pertain. In Malchukov (2006a), this is captured by the Relevance Principle (RelP) which states that a transitivity parameter should be marked on the constituent to which it pertains, that is, the relevant constituent. This principle explains why in most cases A-related parameters are marked on A (see section on differential subject marking below), while O-related parameters pertaining to individuation are marked on O, resulting in DOM. Counterexamples, nevertheless, exist in that sometimes A-features are marked on O or vice versa (Næss 2006; Malchukov 2006a). Consider the following examples from Shipibo-Conibo (Valenzuela 1997), discussed by Kittilä (2002) and Næss (2006). In this Panoan language, A must be marked by the ergative case only when O is referential:

Shipibo-Conibo (Valenzuela 1997)

(20) *e-n-ra* *yapa-ø* *pi-kas-ai*
I-ERG-ASSRT fish-ABS eat-DES-INCOMPL
'I want to eat fish (referential only)'

(21) *ea-ø-ra* *yapa-ø* *pi-kas-ai*
I-ABS-ASSRT fish-ABS eat-DES-INCOMPL
'I want to eat fish (referential or non-referential)'

How are case dependency effects captured from an indexing perspective? Næss (2004a), who generally adopts an indexing approach, suggests that a semantically transitive clause should additionally satisfy the condition of Maximal Semantic Distinguishability of its arguments: if they are not maximally distinguished (e.g. the subject is non-volitional, or object is non-affected), the transitive pattern may shift to intransitive. On a further assumption that non-referential Os are less affected, this principle would explain why ERG is lost in (21). Another approach to the mismatch encoding problem would be to upheld the locality principle, and attribute these mismatches to interference of other factors. One possibility would

be to view the switch from ergative to absolutive in (21) as a side effect of object incorporation (note, however, that the object may be referential in the bi-absolutive construction as well). A more general solution would be to attribute this case alternation to economy. On this approach, a markedness/economy principle like PAIP (Malchukov 2006a) would be held responsible for the fact that in an ergative language an O-feature could affect encoding of A rather than the primary absolutive argument.

In general it seems that some patterns of transitivity alternations are more naturally accounted for under the indexing approach, while some others are better explained in terms of distinguishability. In yet other cases both approaches seem to be equally applicable. Thus, classical cases of DOM, discussed above from a perspective of 'local' distinguishability, can be handled from an indexing perspective as well (cf. Næss 2004a). Note that the type of DOM attested in Hindi and Turkish is expected on H&T's approach as a non-specific/indefinite/inanimate O is less individuated (H&T's parameter J) and thus may be non-eligible for ACC marking.[8]

22.7 Differential Subject Marking

Nowadays most authors in both functional-typological and formal approaches agree that the two functions of case are indispensable, although they differ in the ways how these approaches should be integrated with each other. One way, suggested by Næss (2004a; cf. Testelets 1998 for a similar proposal), was discussed above. Another approach is to view these two motivations as two independent principles (or constraints) in interaction. The latter approach is probably most common in typology, but its consequences for the domain of differential case marking have been insufficiently appreciated until recently (de Hoop and Malchukov 2007; Malchukov 2008b). A welcome consequence of the latter approach is that it can explain asymmetries in differential case marking: the fact that DOM is cross-linguistically consistent, while differential subject marking is not (Woolford 2001; de Hoop and Malchukov 2007). This is so because in case of DOM both functions converge: more prominent (animate, definite/specific) subjects should be marked not only because they qualify as better patients, but also because of the need to

[8] DOM governed by definiteness distinctions may be more difficult to handle under the distinguishing approach because definiteness does not contribute to disambiguation as straightforwardly as animacy does (see de Hoop 2006). Definiteness-based DOM is explicitly related to aspect by some authors (Ramchand 1997; Ritter and Rosen 2001, 2005 a.o.) in whose approaches the definite object moves out of VP to a higher projection (AspP or AGR-OP), e.g. in order to check a feature QUANT in Ritter and Rosen (2001). It is not straightforward how this can be extended to animacy-based DOM.

differentiate them from subjects. In case of subjects these two factors are in conflict, as indexing would favour marking of more prominent subjects (which qualify as better Agents), while distinguishability would favour marking of less prominent subjects (e.g. inanimate) which can be otherwise confused with objects. Therefore DSM is (correctly) predicted to be less consistent cross-linguistically than DOM (see de Hoop, this volume, for further discussion).

Consider the pattern of DSM in Fore, which is a mirror image of the DOM pattern in Awtuw. In Fore (Scott 1978), the ergative determiner (Scott's 'delineator') is used only if O is higher than A on the animacy hierarchy (as in (23)), otherwise A remains unmarked (as in (22)):

Fore (Scott 1978: 116):

(22) *Yagaa wá aegúye*
pig man 3SG.hit.3SG
'The man hits (or kills) the pig'

(23) *Yagaa-wama wá aegúye*
pig-ERG man 3SG.hit.3SG
'The pig hits the man'

The use of the case marker is clearly motivated by the need to disambiguate: ERG case appears on non-prominent (non-human) As which are more likely to be construed as objects than as subjects.[9] While Fore (and other Papuan languages with an optional ergative marking) provides an example of DSM related to global distinguishability, many Australian languages provide evidence for a 'local' distinguishability/markedness approach. In these languages, ergative case is lacking on pronouns, in accordance with Silverstein's markedness approach.

On the other hand, in other languages we find DSM related to the indexing function. In these languages, ERG case is related to agentivity and its loss is symptomatic of a reduction of agentive properties. Consider an example, where ERG switches to an oblique if the subject is non-volitional:

Agul (Ganenkov et al. 2006)

(24) *baw.a neK aTuzu-ne.*
mother:ERG milk:ABS pour_out-PRF
'Mother poured out the milk.'

(25) *baw.a-f-as neK aTuzu-ne.*
mother-AD-ELA milk:ABS pour_out-PRF
'(It so happened that) Mother accidentally spilled the milk.'

[9] See however, Donohue (2005a) for an alternative interpretation.

In the Agul 'involuntary agent construction' (see Kittilä, this volume, for a parallel example from Lezgian), DSM is of the symmetric type and hence it is expected to follow the indexing strategy.

Another consequence of the approach recognizing indexing and distinguishability as two competing motivations (alias, conflicting constraints) is that it accounts for the correlation between differential case marking asymmetries in languages with different alignment types. As noted above, DOM is more typical for accusative languages than for ergative languages. This is due to the fact that DOM in an accusative language does not constitute an Economy (PAIP) violation (the construction retains its primary unmarked argument). In a similar fashion, it is easier to find clear cases for DSM in ergative languages than in accusative. Indeed, Icelandic is exceptional among accusative languages in providing robust evidence for oblique subjects. The experiencer argument within the corresponding constructions in German (of the type *Mir gefällt das Buch*) does not qualify as an oblique subject involved in a DSM pattern, as it does not pass the standard subjecthood tests (see Zaenen et al. 1985 for a classical discussion of Icelandic/German contrasts; cf. also Bayer 2004). Ergative languages, on the other hand, provide better examples for DSM. This pertains both to DSM of the split type (as found in split ergative languages of Australia), and of the global type (as found in Papuan languages like Fore). Also in Daghestanian languages, the ergative and oblique case involved in a DSM pattern do not show a clear contrast with respect to subject properties (see Ganenkov et al. 2006 for Agul).[10]

22.8 Differential Case Marking beyond Core Actants

It is less conventional to speak of differential case marking beyond core arguments even though similar phenomena are found in this domain. For example, Kittilä (2005a) discusses the phenomenon of differential goal marking, referring to languages where animate and inanimate goals are marked differently (for example

[10] Balthasar Bickel (2004b and p.c.) observes that 'oblique subjects' show a considerable variation across languages with respect to syntactic subject properties. Thus, in Tibetan languages oblique subjects usually display syntactic subject ('pivot') properties, while Indo-Aryan languages do it far less consistently, and are more similar to European languages in that respect. This is consistent with our approach insofar as Tibetan languages are consistently ergative, and thus allow for DSM more freely than accusative languages of Europe, while split ergative Indo-Aryan languages fall in between in that respect.

in Finnish animate Goals take the allative, while inanimate ones take the illative case). Similarly, one can speak of differential possessor marking for cases where like in Russian, a special possessive form in *-in* is reserved for proper nouns and kin terms (*Mash-in* 'Masha's', *mam-in* 'mother's') which does not occur on inanimates. Further, a situation where case marking is exceptionally absent on place names may be appropriately characterized as differential locative marking; cf. Persian *ræft-ænd (be) lændæn* [went-3PL (to) London] 'They went to London' where the preposition *be* is optional. This situation is indeed most common with locatives (see Creissels, Chapter 42, and also Ogawa, Chapter 54, on 'case-drop' in Japanese), but may extend to other obliques as well. Consider a situation in Kwaza (a Bolivian isolate; Van der Voort 2004), where both locative *=na* and instrumental *=ko* can be omitted if understood from the context:

Kwaza (Van der Voort 2004: 123)

(26) *awe-mũ(=ko)* *haha-nỹ-ʻdjo-xa-re*
rainwater(=INS) wash-REF-CL:foot-2-INT
'You wash your foot with water?'

Similar examples of 'differential oblique marking' are found in other Amazonian languages. For example, in Tariana, locative, instrumental, and comitative markers may be omitted under similar conditions (Aikhenvald 2003: 185).[11]

An important point here is that the motivations for differential case marking beyond core arguments are basically the same as those discussed for the domain of DOM and DSM (see also Aristar 1996, 1997). Thus alternations in encoding of animate and inanimate goals are more adequately analysed under the indexing approach: namely animate and inanimate goals can be conceived as different roles (say Goals vs. Recipients), which would motivate their differential marking. Differential marking of Possessors is ultimately related to prominence (hence its interaction with the animacy hierarchy). This is underlined by the data from languages like Turkish where both the ACC–NOM alternation and the GEN–NOM alternation on possessors basically refer to the same feature of specificity (cf. Kornfilt 2006 for a discussion of differential case marking in Turkish). Finally, absence of locative marking on place names and other cases of 'differential oblique marking' is a matter of recoverability of a certain meaning, which is related to distinguishability in a broad sense. The marking may be absent if the case-relation is recoverable from the context (from the property of a noun), but must be present otherwise.

[11] No concomitant syntactic changes have been observed for these languages that would suggest that an oblique is 'promoted' to object in these constructions.

22.9 Conclusions

In this chapter we have discussed different types of differential case marking which differ in the grammatical function involved (differential subject and object marking), formal encoding (symmetrical vs. asymmetrical), and lexical restrictions involved (split vs. fluid patterns). It has been argued that some of these patterns are better understood from the indexing perspective while some others follow from the distinguishing/markedness approach. It has further been shown how interaction between these functional contraints can account for certain asymmetries in differential case marking as well as correlations between differential case marking patterns and alignment systems.

Acknowledgement

We gratefully acknowledge the financial support of the Netherlands Organisation for Scientific Research (NWO) for the research reported here (grant no. 220-70-003 'Case cross-linguistically').

CHAPTER 23

CASE AND THE TYPOLOGY OF TRANSITIVITY

SEPPO KITTILÄ

23.1 Introduction

Canonical transitive events involve a volitional and controlling agent and a thoroughly affected patient (see e.g. Hopper and Thompson 1980: 252). Any deviation from this prototype may result in a change in the coding of the denoted event. Case plays a central role in this process: Accusative marking of Patients (NOM–ACC-languages) and ergative marking of Agents (ABS–ERG-languages) are usually associated with the coding of prototypical transitive events, while other case frames (such as NOM–DAT/INS or ABS–DAT/INS) usually code events with a decreased degree of transitivity. The changes in the case marking of (core) arguments may be motivated basically in two ways. First, the changes may follow from verbal morphology as is the case with such derived constructions as passive and antipassive (see Shibatani, Chapter 21). Second, the changes may follow independently of verb morphology. The latter type constitutes the topic of this chapter, because only by excluding the changes in verb morphology can we state for sure that the change in case marking is indeed responsible for the attested difference in transitivity. The focus lies on transitive clauses and the changes in the marking of Agent and Patient arguments. I will first discuss the relation between case marking and transitivity from a formal

978–0–19–920647–6 24-Spencer-ch23 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 357 of 366 May 16, 2008 18:58

perspective, which is followed by a brief discussion of the semantic motivation of the changes.

23.2 On the encoding of transitivity alternations through case

23.2.1 Preliminaries

Different cases (or case frames) are related to different verbs and thus to expressing different event types and for coding the participants involved in these events. Nominative–accusative and ergative–absolutive case frames are associated with the expression of canonical transitive events cross-linguistically (see e.g. Tsunoda 1985: 391). In other words, accusative and ergative cases can be regarded as markers of high transitivity depending on the alignment type of a given language. On the other hand, other cases, most notably dative, instrumental, and different locative cases, usually indicate a lower degree of transitivity, when they occur in clauses with two (pro)nominal constituents. In all the cases in which case marking signals changes in transitivity, arguments bearing overt case marking (such as accusative, ergative, dative, instrumental) are primarily responsible for the expression of (changes in) transitivity. Arguments bearing zero case marking are less relevant in this regard, because in most languages one of the arguments of a two-argument clause must bear the same marking as the only argument of an intransitive clause and their form is not sensitive to transitivity. This means that the zero-marked argument must always be present. Consequently, the focus of my presentation lies on the overtly case-marked arguments.

In what follows, I will illustrate basic transitivity alternations and their relation to case from a formal perspective (the illustration is largely based on Kittilä 2002). In the presentation below, X refers to the single argument of an intransitive clause (which usually, yet not always, bears zero marking), while Y refers to the marked (core) argument of a transitive clause irrespective of the semantic role it bears (it may be agent or patient). The label Z, in turn, covers any argument of a two-argument construction which bears non-core marking. In other words, Z refers to arguments which occur in any other case than nominative, absolutive, accusative, or ergative. I will start discussing transitivity-increasing alternations, after which I will move on to discussing alternations that decrease the transitivity of the affected event. Most of the alternations examined below affect the semantic transitivity of the denoted event, which means that such features as affectedness or agency are

modified. Only in some cases, the effect is pragmatic (Differential Object Marking is the most evident example of this). This seems to be typical of transitivity alternations expressed solely by changing the case.

23.2.2 Transitivizing alternations

This subsection examines alternations which increase the (formal) transitivity of clauses. Three major types of transitivizing alternations can be distinguished based on the number and marking of arguments in the underlying clause and the marking of the introduced arguments. These are discussed in turn below.

(i) X V → X Y V

The first type comprises changes which introduce new arguments into clauses. The type is very frequent and it covers the typical changes between basic intransitive and transitive clauses (changes signalled by verb morphology are excluded here, because the focus of the presentation lies on case). Two examples are found in (1) and (2):[1]

(1) Lezgian (Haspelmath 1993a: 289)
   a. *Get'e xa-na*
      pot break-AOR
      'The pot broke'
   b. *Zamira.di get'e xa-na*
      Zamira.ERG pot break-AOR
      'Zamira broke the pot'

(2) German
   a. *Der Krug zerbrach*
      the.NOM pot break.PAST
      'The pot broke'
   b. *Er zerbrach den Krug*
      he.NOM break.PAST the.ACC pot
      'He broke the pot'

Examples in (1a) and (2a) illustrate typical intransitive constructions of Lezgian and German, respectively. The X argument of these constructions bears zero (nominative or absolutive) marking. (1b) and (2b), on the other hand, exemplify the canonical transitive construction of Lezgian and German. The introduced argument (Y) bears ergative or accusative coding depending on the basic alignment of

[1] It should be noted that the direction of derivation is not necessarily obvious with labile verbs, but for the sake of argument (1) and (2) are here regarded as instances of transitivizing alternations.

978–0–19–920647–6 24-Spencer-ch23 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 359 of 366 May 16, 2008 18:58

the language in question. It is the marking of this argument which is responsible for the morphological transitivity of (1b) and (2b) as either the patient (Lezgian) or the agent (German) is marked as X. The mere number of arguments is not a sufficient criterion of high transitivity here, but the resulting construction needs to be either ERG–ABS or NOM–ACC, because these frames are related to expressing high transitivity in the languages above. For example, as shown in (8) below the Agent can also occur in the adelative in Lezgian, but with an evident difference in semantic transitivity.

(ii) X V → X Z V

The second type comprises cases in which the introduced argument is not a core argument, but rather a part of the clause periphery. Two examples are found in (3) and (4):

(3) Finnish
  a. *Kuppi puto-si*
  cup.NOM drop-3SG.PAST
  'A/the cup dropped'
  b. *Minu-lta puto-si kuppi*
  1SG-ABL drop-3SG.PAST cup.NOM
  'I (accidentally) dropped a/the cup'

(4) Ura (Crowley 1998b: 28f)
  a. *Yarvin y-omrok*
  woman 3SG.DISTPST-steal
  'the woman stole'
  b. *Yarvin y-omrok ŋi namas*
  woman 3SG.DISTPST-steal OBL clothes
  'The woman stole the clothes'

The (a)-examples in (3) and (4) illustrate intransitive constructions in Finnish and Ura. The (b)-examples, in turn, represent extended intransitives derived from (a) (the term 'extended intransitive' is here used in the sense of Dixon and Aikhenvald 2000). In (3b), the Agent occurs in the ablative instead of the (expected) nominative and the Patient retains the nominative marking it has in (3a). In (4b), the Patient is preceded by an oblique preposition instead of occurring in the zero-marked (nominative) form. Although the number of arguments increases in (3) and (4), this does not produce a canonical transitive construction. Instead, these alternations should be classified as (semi-) transitivizing X Z V-alternations, because the arguments introduced to (3b) and (4b) can only occur in an oblique form. The verb would need to be transitivized (causativized/applicativized) for (3b) and (4b) to occur with a canonical transitive frame. It is important to note that the verbs

in (3) and (4) are not derived intransitive verbs, and therefore cannot be taken to represent 'antipassives'. The oblique form of the added argument mirrors the lower degree of transitivity associated with the denoted event. For example, in Finnish the Agent would need to occur in the nominative in the case of highly transitive events.

(iii) X Z V → X Y V

In (1)–(4) alternations have been examined, in which the number of constituents increases. There are also alternations which transitivize a clause without affecting the number of arguments. Two examples of this are provided in (5) and (6):

(5) Shipibo-Conibo (Valenzuela 1997: 197)
   a. *Ea-Ø-ra* *yapa-Ø* *pi-kas-ai*
      1-ABS-ASSRT fish-ABS eat-DESD-INC
      'I want to eat fish' (referential or non-referential)
   b. *E-n-ra* *yapa-Ø* *pi-kas-ai*
      1-ERG-ASSRT fish-ABS eat-DESD-INC
      'I want to eat the fish' (referential only)

(6) Finnish
   a. *Mies* *rakas-ti* *koira-a-nsa*
      man.NOM love-3SG.PAST dog-PART-3POSS
      'The man loved his dog'
   b. *Mies* *rakas-ti* *koira-n-sa* *kuoliaaksi*
      man.NOM love-3SG.PAST dog-ACC-3POSS to.death
      'The man loved his dog to death'

Examples in (a) illustrate *XZV*-constructions (extended intransitives) of Shipibo-Conibo and Finnish. In (a) either the Agent or the Patient occurs in a form not typical of the canonical transitive construction of the languages in question. In Shipibo-Conibo, both arguments of a desiderative construction, as in (5a), occur in the absolutive case (the Agent occurs in the ergative in a canonical transitive construction). The low transitive marking of (5a) follows from the irrealis nature of the event it denotes. Desiderative events, by definition, potentially occur in the future. The difference between (5a) and (5b) lies in the fact that the patient of (5b) is referential. As generally known, definite (animate) patients receive a more elaborate formal treatment in a number of languages, a phenomenon known as Differential Object Marking (see Bossong 1985a for DOM in general and Næss 2004b for an elaborate discussion of DOM from the view point of affectedness). The formal difference between (5a) and (5b) is governed by the same principle, but with the difference that the marking of the Agent is sensitive to the definiteness (referentiality) of the Patient, while the marking of the Patient remains unchanged.

978–0–19–920647–6 24-Spencer-ch23 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 361 of 366 May 16, 2008 18:58

In Finnish, the verb *rakastaa* 'love' typically takes an object in the partitive case, as in (6a). The partitive marking of the Patient underlines the fact that 'love' is not a dynamic event that would have a salient effect on its target. In (6b), on the other hand, the event 'love' is seen as having a direct effect on the patient, whence the Patient bears accusative coding.

23.2.3 Intransitivizing alternations

Thus far transitivity-increasing alternations have been examined. It is typical of these alternations that an accusatively or an ergatively marked argument occurs in a clause it was not a part of previously. It either replaces a non-core argument or it is introduced as a completely new argument. The argument in the accusative or in the ergative is then responsible for expressing the high transitivity of the denoted event. In addition to the transitivity-increasing alternations, there are also transitivity-decreasing alternations, whose function is the exact opposite of the alternations examined previously in that they eliminate an accusatively or an ergatively marked argument. These constitute the topic of this subsection.

(i) X Y V → X V

The first and the least controversial type of transitivity-decreasing alternation is illustrated by cases in which one of the arguments of a transitive clause is completely eliminated. This elimination typically has the consequence that the only overt marker of transitivity gets omitted as well, because the only remaining argument is not marked overtly for case. An example is found in (7):

(7) Finnish
Mies *sö-i* *(leivä-n)*
man.NOM eat.PAST-3SG (bread-ACC)
'The man ate (the bread)'

This is an example of what has been labelled as Indefinite Object Deletion (IOD) (see Næss 2004a: Chapter 3 for a more detailed discussion of IOD and the references therein). In IOD an object, which is indefinite or whose identity is not important, is completely eliminated. In NOM–ACC-languages such as Finnish, this has the consequence that the argument that encodes transitivity overtly is deleted. There are no other changes in the clause structure. In (7), the overtly case-marked argument is omitted from the clause and we are left with a zero case-marked argument. There is thus no overt transitivity marking in (7).

(ii) X Y V → X Z V

This is the mirror image of X Z V → X Y Z discussed in the previous section and is illustrated in (8) and (9):

978–0–19–920647–6 24-Spencer-ch23 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 362 of 366 May 16, 2008 18:58

(8) Lezgian (Haspelmath 1993: 289, 292)
a. *Zamira.di get'e xa-na*
Zamira.ERG pot break-AOR
'Zamira broke the pot'
b. *Zamara.di-waj get'e xa-na*
Zamira-ADEL pot break-AOR
'Zamira broke the pot accidentally/involuntarily'

(9) Lhasa Tibetan (examples courtesy of Scott DeLancey)
a. *Sta=re-s shing(*-la) 'chad-pa*
axe-ERG tree cut
'Cut the tree with an axe'
b. *Shing-la sta=re-s gzhus-pa*
tree-LOC axe-ERG hit
'Hit the tree with an axe'

In (8), the marking of the Agent varies according to the overall transitivity of the denoted event. The Agent occurs in the ergative in the case of canonical transitive events, while the marking shifts to the adelative in case the agency and thus the transitivity of the denoted event is lowered. In Lhasa Tibetan, the marking of the Patient is sensitive to affectedness. The (zero) absolutive marking (and thus the ERG–ABS case frame) is related to high degrees of affectedness (and transitivity), while overt (locative) marking (and the ERG–LOC case frame) is associated with lower transitivity (for a more detailed discussion of case frames and transitivity, see e.g. Tsunoda 1981b and Malchukov 2005). The change is from zero to overt marking in (9).

23.3 Semantics of case in transitivity alternations

In Section 23.2, I examined the contribution of case to the expression of transitivity from a primarily formal perspective. In this section, the semantics of case in the expression of transitivity will be examined. I will discuss the motivation behind the attested changes, the features different cases are related to, and I will also briefly discuss the difference between grammatical and semantic cases in the expression of transitivity.

We can distinguish between two types of transitivity alternations according to their motivation. In addition the contribution of case to these alternations varies. First, there are alternations which follow from the inherent semantics of verbs

978–0–19–920647–6 24-Spencer-ch23 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 363 of 366 May 16, 2008 18:58

used for describing events. Examples of this are provided in (1), (2), and (9). In these cases, changes in the case marking of arguments rather mirror the differences in verb semantics instead of making any significant contribution of their own. Different verbs govern different case frames (see again Tsunoda 1981b and Malchukov 2005). The second type of alternation is illustrated by cases in which the case marking is primarily responsible for coding a change in transitivity. Examples are found in (6) and (8). In these cases, the inherent verb semantics cannot be responsible for the attested changes, because the verb is the same in both cases. This means that the changes in the reading of the clause can only follow from the change in case marking, which is the primary indicator of transitivity here. For example, in (8) from Lezgian, the use of the ergative (which is the default case for basic transitive clauses) implies high agency, while the use of the adelative indicates that the degree of agency associated with the denoted event is lower than expected. The event as such, along with its salient result, remains the same in these cases.

Different cases can be classified depending on the semantic feature of transitivity they encode. For example, the ergative case is typically (yet not necessarily) associated with the coding of agency (see (8) from Lezgian), while the accusative case usually encodes affectedness (see (6) from Finnish). Ergative and accusative canonically code the two major features of high transitivity. Moreover, because they are the only overtly marked cases of basic transitive clauses, they are also primarily responsible for coding the high transitivity of the event in question. The two patterns differ from each other only in the overtly marked feature of transitivity. The nominative and absolutive cases can be considered as default, zero-marked, cases of basic transitive clauses and their contribution to transitivity coding is thus less relevant in many cases (see, however (9) from Lhasa Tibetan).

As said above, affectedness and agency are the two most important single features of linguistic transitivity (see also e.g. Hopper and Thompson 1980, Kittilä 2002, and Næss 2004a among many others). This also applies to expressing transitivity through case. Other features of high transitivity that make a contribution to the coding of events include aspect and tense as shown in (10) and (11):

(10) Hindi (Mohanan 1994: 70)
  a. *Raam-ne ravii-ko piitaa*
     Ram-ERG Ravi-ACC beat-PRF
     'Ram beat Ravi'
  b. *Raam ravii-ko piittaa hai*
     Ram.NOM Ravi-ACC beat-IMPF be-PRES
     'Ram beats Ravi'

(11) Samoan (Mosel and Hovdhaugen 1992: 111)
  a. *Sā faitau (ʻuma) e ulika le tusi*
     PAST read (all) ERG Ulika ART letter
     'Ulika read the (whole) letter'

b. *Sā faitau ulika i l=a=na tusi*
PAST read Ulika LD ART=POSS=3SG letter
'Ulika read her letter' (lit. Ulika read in her letter)

The examples from Hindi illustrate a situation in which tense determines the marking of arguments (Agent) and thus makes a contribution to the transitivity of clauses as well. In (11) from Samoan, aspect conditions the case marking of arguments: the case frame is ERG–ABS in perfective aspect, while LOC–ABS signals imperfective aspect. The relation of tense and aspect to transitivity is rather transparent, which makes it easy to see why these features make a contribution to the formal transitivity of clauses. For example, only events in the past tense have a salient result and a thoroughly affected patient, and the patients of successfully completed events (perfective aspect) are more thoroughly affected than patients of on-going events (imperfective aspect). This kind of tense/aspect-conditioned split in case marking is very typical cross-linguistically.

Accusative and ergative, along with the zero-marked nominative and absolutive cases, are usually labelled as structural/grammatical cases, because they code grammatical functions (of Agent and Patient typically, see e.g. Blake 1994: 32–5). In addition, they can also be regarded as the 'highly transitive cases', because they appear in clauses denoting highly transitive events, as was noted above. Other cases, such as instrumental, dative, locative, and genitive, in turn, are usually labelled as semantic cases, which means that they are more intimately related to the expression of semantic content (dative is intermediate between grammatical and semantic cases, because it also has grammatical functions, such as marking the indirect object relation). As for coding transitivity, semantic cases usually appear in clauses denoting less than transitive events. Examples are found in (3) and (4), for example. Even though semantic cases are usually more intimately associated with the expression of semantic content (such as location or instrument), it is harder to associate them directly with the expression of any specific transitivity feature. For example, the instrumental case may be used for coding the agent of passive, the patient of antipassive, less affected patients, and less agentive agents across languages. The dative also expresses a similar array of functions, in addition to which the dative often encodes the experiencer of experiencer constructions (such as '*a person* loves an individual'). The only feature these cases share is decreased transitivity. We can thus make the generalization that grammatical cases are associated with high transitivity, while semantic cases code decreased transitivity.

What is also noteworthy with regard to the relation between transitivity and case is that irrespective of the basic alignment type of a given language, the overtly marked arguments (be they marked by grammatical or semantic cases) are primarily responsible for the expression of transitivity. An illustrative example of this is found in (8) from Lezgian, where the ergative (grammatical case) codes events

978–0–19–920647–6 24-Spencer-ch23 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 365 of 366 May 16, 2008 18:58

instigated by a volitionally acting agent, while the adelative (semantic case) appears in clauses denoting involuntarily instigated events. The unmarked cases, on the other hand, are unmarked also as regards the coding of transitivity. In some cases, such as passive and antipassive, the semantic role borne by the unmarked argument changes (for example, in passive the nominatively marked argument encodes the patient instead of the agent). But here, too, it is the overtly marked argument that signals the decreased transitivity of the clause.

PART IV

CASE IN (PSYCHO)-LINGUISTIC DISCIPLINES

CHAPTER 24

THE ACQUISITION OF CASE

SONIA EISENBEISS
BHUVANA NARASIMHAN
MARIA VOEIKOVA

24.1 Introduction

Case is one of the most heterogeneous nominal morphological categories: the number of case forms in morphological paradigms, the syntactic and semantic functions of case, and the set of declension classes differ even in typologically similar languages. Hence, the acquisition of case presents the child with a major learning challenge. Our survey presents empirical studies and theoretical perspectives on the acquisition of case in children, focusing on generative, natural morphology, cognitive-functional, and usage-based approaches. Our empirical focus will be on the acquisition of accusative, ergative, and split case systems, and our theoretical discussion will concentrate on (i) productivity in children's early case forms, (ii) the role of nature and nurture in the acquisition of case, (iii) form–meaning mappings in the acquisition of case, and (iv) the time course of case development.[1]

[1] For reasons of space, we will not discuss the acquisition of case marking by second language learners. See Hawkins (2001) and White (2003) for overviews and additional references.

24.2 Productivity in children's early use of case forms

In comparing case development within and across languages we need clear criteria to determine whether case forms in early child language are rote-learned, or used productively and in systematic contrast to other forms. For instance, case-marked forms are often found very early in grammatical development, especially in agglutinating languages like Turkish and Finnish (Laalo 2002; Küntay and Slobin 1999, 2002) but we can only say that a child has acquired these forms if a number of criteria are met.

First, the forms are correctly applied in the appropriate contexts suggesting that the child understands the meaning of the form. But it does not tell us whether this form is bound to a narrow range of contexts, or if its use is motivated by adult-like notions like 'agent' or 'subject'. There is some evidence that the meaning of case forms is acquired gradually: peripheral or more abstract meanings are attested later than prototypical meanings (Christofidou and Stephany 1997: 138; Savickienė 2003: 64–90; Stephany and Voeikova 2007). For instance, Russian children use the accusative for names of familiar physical objects (*daj konfetku* 'give-IMP candy-DIMINUTIVE.ACC') before they use it as a time marker (*podoždi minutku* 'wait-IMP minute-DIMINUTIVE.ACC') (Ceytlin 1989: 68 ff.).

Second, the case form contrasts with at least one other form with the same stem (Gvozdev 1961), e.g. Russian *blin-a* 'pancake-SG.GEN' vs. *blin* 'pancake-SG.NOM'. This indicates that the child has started to distinguish between stems and suffixes, suggesting the emergence of case forms.

Third, the case form (e.g. an affix) is used with at least three different stems (Bittner et al. 2000: 4–5). This shows emerging mastery of a morphological pattern that is necessary for productive use (Kilani-Schoch and Dressler 2002: 45–6).

Thus, children's use of a given case form does not necessarily indicate that it has been acquired, but must be interpreted differently depending on the range of contexts, as well as the variety of contrastive combinations in which it occurs. Productivity in children's early uses of devices such as case-marking has important consequences for theoretical debates on how children acquire language. Early productivity raises the issue of whether children are equipped with prelinguistic (innate) biases that enable them to rapidly discern the underlying regularities conditioning the use of surface forms. On the other hand, protracted use of rote-learned forms and gradually evolving productivity would suggest that such knowledge is constructed from scratch. In the next section, we turn to a discussion of approaches that differ in the degree of importance they give to nature versus nurture.

978–0–19–920647–6 25-Spencer-ch24 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 371 of 383 May 16, 2008 18:59

24.3 Nature or Nurture in the Acquisition of Case

The nature of children's early generalizations in case acquisition has been the focus of much debate revolving around the so-called 'logical problem of language acquisition' (Baker 1979, Pinker 1989, Braine 1992, Bertolo 2001). Children only hear a limited sample of their target language, and thus have to generalize over individual input utterances in order to comprehend and produce new utterances. If children's hypothesis space for these generalizations were completely unconstrained, children might make incorrect generalizations and one would have to explain how they would ultimately reject them. For example, in Japanese, the accusative marker *o* of preverbal direct objects is optional. Children acquiring Japanese might make the incorrect generalization that *o*-markers are optional on direct objects in any sentential position. In order to reject this overgeneral hypothesis, children would need to know that *o*-omissions in non-preverbal positions are ungrammatical. But explicit correction of this sort ('negative evidence') is not systematically available to all children at all developmental stages (Brown and Hanlon 1970, Marcus 1993).

Furthermore, one has to explain why certain generalizations are never made. For example, German children hear many prepositional phrases with syncretic nominative/accusative forms that might suggest that prepositions can assign nominative (e.g. *für meine Hühner* 'for my-NOM/ACC.PL chickens'). Moreover, nominative is the most frequent case in German (Meier 1967) and thus a likely candidate for overgeneralizations.[2] Interestingly, nominative is never overgeneralized to arguments of dative-assigning prepositions, for example children do not produce *Spiel mit ich!* 'Play with I' instead of *Spiel mit mir* 'Play with me!'.

Thus, a challenge for developmental psycholinguists is to explain how children can avoid or recover from incorrect generalizations about case-marking without systematic explicit correction. Broadly speaking, various accounts can be viewed as arguing for the relatively greater role of nature or of nurture in guiding children's acquisition process.

24.3.1 Generative approaches

Faced with the 'logical problem of language acquisition', generative linguists postulate that children's language acquisition is guided by an innate Universal Grammar (UG), that specifies the grammatical categories and universal well-formedness constraints for linguistic structures. UG is assumed to restrict children's hypothesis

[2] Overgeneralizations may also occur due to the general Case Preference Principles discussed in Bader and Lamers (Chapter 26); however, we are not aware of any investigation of their role for the early speech production of children at the current time.

space so that they can only make correct generalizations or generalizations that can be rejected without explicit correction. For the acquisition of case, generative linguists have postulated: (i) innate case-categories, (ii) a Case-Filter principle that requires all noun phrases to carry abstract case, (iii) principles that govern case-assignment to noun phrases with particular thematic or syntactic roles, and (iv) principles restricting optional case-marking to particular syntactic environments (see Chomsky 1981, 1986). These UG-components prevent children from hypothesizing (i) that case markers can be freely omitted in any sentential position, (ii) that prepositions can assign nominative, (iii) that case assignment is random or based on factors other than thematic/syntactic role (e.g. phonological factors), etc.

In exploring the role of UG in case acquisition, researchers distinguish between structural versus lexical cases. Structural cases are associated with particular structural positions: nominative with the subject position in accusative languages, ergative with the transitive subject position in ergative languages, etc. If UG-principles constrain such associations between cases and positions (Chomsky 1981), children should be able to use these UG-principles together with their input to determine which types of links their target language exhibits (see section 24.4). And once these links are established, structural case should be used correctly. In contrast, lexical (idiosyncratic) cases are not predictable on the basis of structural positions, but associated with particular case-assigning lexemes. Hence, children must learn these associations lexeme by lexeme. For instance, German children have to learn that *helfen* 'to help' assigns dative case and not the accusative to its direct object. Before they have acquired this lexeme-specific property, they should overgeneralize the structural accusative case to the direct object of *helfen*. And indeed, Russian and German children who produce and comprehend structural cases in a target-like manner frequently overgeneralize structural cases to contexts for lexical case (see Babyonyshev 1993, Eisenbeiss 1994a, 2003, Eisenbeiss et al. 2006; see Lamers and Ruigendijk, Chapter 27, for a discussion of structural and lexical case-marking in aphasia).

But interestingly, children's errors are not restricted to lexical case-marking. Systematic errors for structural cases have been found in children who already exhibit target-like case distinctions. Young English-speaking children often use both target-like nominative subject pronouns and non-nominative pronoun forms that are not appropriate for subjects (e.g. *I/*me go*). In order to account for the apparent violations of UG that are predicted not to occur in child language, proponents of the so-called Agreement-Tense-Omission Model appeal to *maturation*, arguing that the non-nominative subject pronouns occur during a maturational stage in which a prerequisite for structural case-marking (finiteness-marking in main clauses) is not yet obligatory for children (Schuetze and Wexler 1996). During this stage, children may produce utterances lacking the finiteness features required for nominative assignment to subjects. This lack of a case assigner can explain the use of the accusative form, which Schuetze and Wexler take to be an underspecified default case

form that can appear in contexts without overt case assigners, for instance in the subject position of sentences without a finite verb (*what him worry?*). In languages where the underspecified default form is the nominative (e.g. in German), there should be no systematic use of non-nominative subjects even if the verb lacks finiteness features. And indeed, such errors are not documented (see Eisenbeiss 2003, Eisenbeiss et al. 2006 for overviews).

Recent theoretical advances have resulted in some convergence between generative and functionalist (or usage-based) approaches as minimalist generative models try to derive the innate linguistic constraints suggested in earlier research from more general cognitive principles (Chomsky 1995). In particular, various researchers have replaced case-assignment principles (see also Bader and Lamers, Chapter 26) by a general Specificity Principle that gives operations involving specific information precedence over operations with less specific information (Kiparsky 1997, Wunderlich 1997, Eisenbeiss 2003). Such a principle can capture the distribution of case markers because these markers differ with respect to the specificity of the information they encode; for instance, accusative markers on verb arguments are restricted to lower-ranking arguments of transitive verbs. In contrast, nominative can be assigned to single arguments of intransitive verbs or the higher argument of transitive verbs. Hence, they are less restricted in their distribution and thus contain less specific information than accusative markers. General principles like the Specificity Principle have been taken up in accounts of case acquisition that try to bridge the gaps between generative and functionalist or usage-based approaches (see section 24.4).

24.3.2 Functionalist approaches

Proponents of functionalist or usage-based approaches assume that domain-general cognitive and socio-pragmatic principles suffice to constrain children's hypothesis space because children's input provides rich, structured information about the target language. For instance, cross-linguistic research shows that parents present their children with so-called 'variation sets', i.e. sequences of adult utterances with a constant communicative intention and different types of variation in form (Bowerman et al. 2002, Eisenbeiss 2003, Küntay and Slobin 1996, 2002). These variation sets may involve variations of case marking, such as *I* vs. *me* and *she* vs. *her* in the variation set ***I*** *cannot see* ***her***, *but* ***she*** *can see* ***me!*** Hence, variation sets in the input can highlight case contrasts, providing sufficient information for the child to infer the appropriate form–function correspondences in the case system of the target language. Other researchers have shown that at least some parents in Western cultures frequently reformulate their children's non-target-like word forms, producing correct forms which contrast with the child's non-target-like form (see e.g. Chouinard and Clark 2003). Such a contrast could alert the child

to the error and thus provide a type of negative evidence even when the adult does not explicitly say that the child's case-marked form is incorrect. In addition to pragmatic directions from caregivers, children are also exposed to information in the form of distributional patterns of different case-forms in the input. Children are sensitive to such influences in their acquisition of case-marking – both type-frequency and phonological diversity have been shown to influence children's acquisition of case-marking in Polish (Dąbrowska and Szczerbiński 2006).

Input-related factors are also assumed to contribute to children's case-marking errors that are accounted for by maturational factors within generativist approaches. For instance, Tanz (1974) attributes English-speaking children's accusative case overgeneralizations to subjects (see section 24.3.1) to the fact that the accusative forms of personal pronouns appear in more diverse syntactic contexts than nominative forms in the input, and occur more often in the sentence-final position which is relatively salient. Other researchers have stressed the role of morphophonological input patterns. Rispoli (1999) argues that the Agreement-Tense-Omission Model (see section 24.3.1) does not explain why *her* is more likely to occur instead of *she* than *his* or *him* instead of *he*. He attributes this observation to the fact that *her* is used in both possessive and beneficiary function and therefore is more 'loaded' and also more frequent.

Further evidence for the influence of the input is provided by language-specific differences in children's choice of the default form of nouns. Typically, in every declension type one case form is used as default, or base form. In Russian child language, the preferred base form is the nominative which is uninflected in productive masculine nouns and is marked with the *-a* ending in productive feminine nouns (Gvozdev 1961: 136). In Lithuanian, nominative nouns have several different case endings, but instead of selecting one form as a default, young children sometimes prefer to use a protomorpheme *-(i)a*, e.g. *bulia* (bulvė) 'potato' or *palelia* (piniginė) 'purse' (Savickienė 2002: 106).

Other error types suggest that children's preferences for unambiguous, overt forms may lead to case errors. For instance, children are sensitive to case distinctions before they have mastered the different case forms for each declension. A classic study on the acquisition of Russian grammar by Gvozdev (1961) based on diary data from his son Zhenja showed that there was no evident preference for a particular declension class (DC) in Zhenja's speech. Instead, for every case a particular inflectional ending (except zero) is preferred; for example Zhenja erroneously says **kresl-U* 'armchair-ACC' (DC 1) instead of *kresl-a* (DC 2) but **sobak-OV* 'dog-GEN.PL' (DC 2) instead of *sobak-0* (DC 1). Slobin (1966) developed Gvozdev's ideas and attributed such errors to 'inflectional imperialism': children initially choose correct case forms, but prefer a certain inflectional ending (irrespective of DC) for every case except the nominative. Which endings are selected is determined by preferences for unambiguous and overt, i.e. non-zero, endings and the avoidance of exceptional grammatical forms.

978–0–19–920647–6 25-Spencer-ch24 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 375 of 383 May 16, 2008 18:59

While these studies suggest that language-specific morphological properties and children's own preferences may lead to case errors, recent research shows that (i) the percentage of overgeneralizations in children's speech is comparatively low and (ii) overgeneralizations mostly occur in a limited period of time. In particular, children acquiring Slavic and Baltic languages, such as Croatian, Lithuanian, Russian, exhibit low overgeneralization rates (Jelaska, Kovačević and Anđel 2002: 180, Savickienė 2002: 110–12; Voeikova and Gagarina 2002: 131–3). Whereas generative acquisition researchers attribute these findings to the early availability of constraints on case-marking, proponents of usage-based accounts view these findings as evidence for a strong reliance on rote-learned forms and small-scale analogies in early grammatical development.

In sum, empirical studies have found systematic patterns of case-marking (errors) that may be explained in terms of innate constraints or input patterns coupled with general cognitive mechanisms of generalization and analogy formation. We now turn from our discussion of children's acquisition of case forms to examine more closely the issue of how children determine their functions.

24.4 Acquiring Form–Meaning Mappings

Independently of their views on the logical problem of language acquisition, researchers must explain how young children link language-specific case forms in the input with their functions. Prior research shows that children are remarkably sensitive to the morphosyntactic devices used in the input to mark semantic relations, such as word order in English vs. case-marking in Turkish (Slobin 1982). In fact, despite the importance of canonical ordering in child language (Braine 1976, Pinker 1984), children learning languages that use morphological case markers to indicate semantic relations, e.g. Turkish, comprehend agent–patient relations earlier than children learning a language such as English that uses word order (Slobin 1982). Slobin claims that the developmental advantage of 'local cues' such as case markers lies in the fact that they have phonological content and apply to particular nouns, whereas word order has no phonological content and constitutes a 'global cue' requiring that the entire sentence be taken into account in inferring agent–patient relations.

Cross-linguistic research suggests that other factors that play a role in the acquisition of agent–patient marking include the frequency and consistency with which different forms are used in the input, as well as children's cognitive and linguistic skills at a given point in development (Bates and MacWhinney 1989). For instance, young children acquiring case-marking languages such as Hungarian or

Serbo-Croatian are inclined to use animacy rather than case-marking as a cue to agent–patient relations when asked to act out sentences that provided both types of cues (Bates and MacWhinney 1989).

Whereas children's experience with frequently used forms in the input and recurrent situations in their lives are important factors influencing their comprehension and encoding of agent–patient relations, it has been suggested that children also start out with certain biases in their use of grammatical markers to convey 'who did what to whom'. For instance, Slobin (1985) argues that children initially underextend grammatical marking to certain event types, e.g. manipulative activity scenes involving basic causal events in which an agent carries out a physical and perceptible change of state in a patient by means of direct body contact or with an instrument under the agent's control. Children use the grammatical markers they hear in the input (accusative or ergative inflections, word-order patterns) to mark aspects of such scenes. This claim is supported by the finding that ergative inflection first appears only on subjects of verbs such as *give*, *grab*, and *hit*, in Kaluli; and tends to be omitted in sentences with verbs such as *say* and *see* (Schieffelin 1985). Children acquiring English also make pronoun case errors that reflect a strong semantic motivation to indicate a distinction showing the degree of subject involvement in the action expressed (Budwig 1989). Verbs expressing actions which are overtly agentive and require some degree of control on the part of the (first person) subject are used with a subject pronoun *my*, while those which were not expressive of agentivity were marked with a subject pronoun *I* (e.g. *I like cookies*).

But in many cases, children learning different languages do not exhibit the pattern of case-marking errors that are predicted if they start out with semantic notions such as agent. For instance, in ergative languages, they do not overextend 'ergative' case marking to agentive participants of intransitive actions (Bowerman 1985, Ochs and Schieffelin 1995, Pye 1990), and the ergative marker in split-ergative languages such as Hindi is not incorrectly extended to all transitive agents (Narasimhan 2005). Moreover, more recent studies on languages like German could not find any initial restrictions of case-marking to particular verb classes (Eisenbeiss 2003). This suggests that children's early semantic categories may be shaped by language-specific categories in the input (Bowerman 1985). An alternative possibility is that the particular semantic notions (agent) that researchers have posited as guiding early grammatical development are not the right ones but may be considerably more abstract (Eisenbeiss 2003).

A second important issue in this context has to do with *(dis)continuity*. For instance, some researchers assume that children's early grammars can be characterized in terms of semantic, e.g. 'agent', 'action', rather than formal grammatical categories, e.g. 'subject', 'noun' (Braine 1992; Schlesinger 1982). If children start out with the assumption that 'nominative' case is linked with 'agent', what prompts children at a later stage to replace 'agent' with adult-like syntactic notions such as 'subject'? Some researchers propose that children initially start out with semantic

underpinnings for grammatical markings, but use distributional evidence to gradually expand their semantically based categories to resemble the formal grammatical categories of adult grammar (e.g. Braine 1976, Maratsos and Chalkley 1980). For example, children might use nominative only for agents, but may notice that nominative is also used for event participants that are not strictly agentive (e.g. subjects of *see*). Thus, a clear distinction between functional and positional processes in speech production (see Chapter 25 by Melinger, Pechmann, and Pappert, on the evidence of such separation) seems not to be present at the early stages of language acquisition.

In contrast, generative researchers suggest that even children's earliest case systems are characterized by the formal grammatical categories of adult grammar (Pinker 1984). Proponents of 'semantic bootstrapping' approaches claim that children use innate form–meaning links to identify case forms in the input and then compare transitive and intransitive sentences to classify these forms. For instance, Pinker (1984) argues that children possess innate links between the patient role of transitive verbs and the case categories 'accusative' and 'absolutive' – and can thus find accusative or absolutive markers by looking for patient noun phrases in the input (see Pinker 1989 for a slightly revised model). Once markers have been identified as 'accusative' or 'absolutive', children must compare intransitive agents to transitive agents and patients to find out whether their target language has an accusative system (intransitive agent=transitive agent) or an ergative system (intransitive agent=patient; see Bowerman 1985).

However, children do not seem to start out with case markers for either transitive or intransitive verbs and then extend them. Rather, they acquire markers for both verb types at the same time (see Morikawa 1989 for Japanese, Eisenbeiss 2003 for German). Moreover, comparing transitive and intransitive verbs would require an unambiguous distinction between these two verb types; but in a language with argument ellipsis (e.g. Japanese), transitive verbs frequently occur with just one argument or no arguments at all (Rispoli 1991). Thus, the semantic bootstrapping account would have to be extended to explain how children nevertheless manage to distinguish transitive and intransitive verbs, possibly by making use of discourse information to detect ellipted arguments (Ratitamkul et al. 2004, Narasimhan et al. 2005).

In sum, models without innate case categories have difficulty explaining how children acquire a formal system from a meaning-based system. On the other hand, models with innate principles and grammatical categories have problems in capturing children's generalizations and cross-linguistic variability in case-marking. Therefore, some researchers have suggested alternative approaches, according to which (i) case development is driven by general cognitive constraints and mechanisms for analysing distributional patterns in the input (Eisenbeiss 2003), and (ii) children gradually build up an inventory of stored form/meaning-mappings (Tomasello 2003).

The lexicalist structure-building approach advocated by Eisenbeiss (2003) is based on features that encode relationships between arguments (Kiparsky 1997, Wunderlich 1997; see Primus, Chapter 17). For instance, lower-ranking arguments like transitive patients can be characterized by the feature [+hr] (=there is a higher argument present). This feature distinguishes them from other arguments (e.g. from the sole agent-argument of intransitive verbs or a higher-ranking agent-argument of transitive verbs). The feature [+hr] can also be used to characterize accusative forms because they mark lower-ranking arguments. Children have to find out which input forms carry case features like [+hr]. This process is driven by the Specificity Principle that gives operations involving specific information precedence over operations with less specific information (see section 24.3). This principle is only applicable if the use of contrasting case forms involves different feature specifications. Therefore, case contrasts in the input should prompt children to search for distinctions that correlate with the form contrast. In this search, children use their lexical knowledge to determine which event participants noun phrases like *Jack* or *John* refer to in sentences like *Jack tickles John.* Then, children can use their conceptual knowledge to determine the causal relations between these event participants. These causal relations can then be mapped onto asymmetric relations between semantic arguments in hierarchical representations. The mapping is constrained by the general cognitive Relation-Preservation Principle that requires homomorphic mappings of asymmetric relations. For instance, if an event participant *Jack* precedes another participant *John* in the causal chain, the Relation-Preservation Principle requires that the semantic argument referring to *Jack* has a higher position than the argument corresponding to *John*. Children can then correlate the relative hierarchical position of semantic arguments with the corresponding morphological marking. For instance, they can find out that accusative is restricted to lower-ranking arguments. Based on such correlations, positive feature specifications can be integrated into lexical entries for case forms (e.g. [+hr] for accusative). Later, children generalize over individual case-marked word forms and incrementally build up lexical entries for case affixes and irregularly inflected case-marked forms.

A different approach has been provided by Tomasello, Lieven and colleagues (see Tomasello 2003). They argue that children start out with limited generalizations centred around individual lexical items and phrases that they gradually extend by analogy. The *continuity* problem – 'how to get from here to there' – is circumvented by assuming that children's grammars are not qualitatively different from adult grammars in the first place. But whereas generativist accounts assume that children's grammars are characterized by the adult-like formal, abstract categories, Tomasello's approach makes the opposite assumption. Relying on cognitive-functional approaches to grammar such as Construction Grammar (Fillmore et al. 1988, Goldberg 1995), Tomasello suggests that adult grammars consist of interrelated schemas that are language-specific, and are characterized by various degrees

of abstractness, ranging from idioms with concrete lexical items (e.g. *kick the bucket* 'die'), to abstract templates characterized by grammatical roles such as subject–predicate. Children start out with 'verb-islands' (Tomasello 1992) consisting of specific verbs and their arguments, e.g. 'hitter-hit-hittee'. In such an approach, morphosyntactic devices such as case-marking or word order are initially associated with individual verbs (and (pro)nouns). Only when a critical mass of verb-islands is acquired do children make generalizations such as associating agent or subject with preverbal position in languages such as English (Akhtar and Tomasello 1997). Frameworks such as Role and Reference Grammar that motivate case-marking phenomena in semantico-pragmatic terms in adult grammar (Van Valin, Chapter 7) can also provide a way to circumvent the continuity problem since no sharp discontinuity needs to be posited between early meaning-based systems and case-marking systems in adults.

In sum, faced with the logical problem and the problems accounting for discontinuities, both generative and non-generative researchers have recently proposed incremental acquisition mechanisms and models of semantic and grammatical representations that are based on stored form/meaning units. Proponents of these models claim to capture children's generalization patterns and cross-linguistic variability better than purely syntactic or semantic models.

24.5 The Time Course of Case Development

Any acquisition model must explain why case acquisition takes several years and case-marker elements are initially absent or optional. Faced with this 'developmental' problem, some generative researchers argue that the UG-principles which constrain case marking become available only later in development, due to the maturation of the neural system that underlies language functions (Felix 1987, 1992) (see section 24.3.1). However, even the earliest structures that children produce do not contain the random errors one would expect if UG-constraints were unavailable (Crain 1991). Therefore, most generative psycholinguists assume that UG-principles are available from the onset of linguistic development (Pinker 1984).

However, they disagree about the nature of categorization predispositions and grammatical representations in young children. According to continuity approaches, categorization predispositions are available at all developmental stages (Pinker 1984), whereas proponents of maturation claim that case and other grammatical categories only 'mature' around children's second birthday (e.g. Radford

1990). With respect to children's grammatical representations, full-competence approaches assume early adult-like representations (e.g. Schuetze 1996), while proponents of structure-building approaches argue that not all case features are syntactically active in the two-word stage, either because of the late maturation of case categories (Radford 1990) or due to a lack of knowledge about the case forms of the target (e.g. Clahsen et al. 1994, 1996, Eisenbeiss 1994a, 2000, 2003).

Given a maturational approach, one might expect that all case markers are acquired at similar stages of linguistic development. However, numerous studies have shown developmental dissociations. For example, dative markers appear later than the nominative/accusative distinction in languages like Japanese or German or the absolutive/ergative distinction in languages like Basque (see Clahsen et al. 1994, Eisenbeiss 2003, Eisenbeiss et al. 2006 for overviews; see Bader and Lamers, Chapter 26, as well as Lamers and Ruigendijk, Chapter 27, for a discussion of the marked nature of dative case and its implications for language comprehension and language impairment in aphasia). And the nominative/accusative distinction in German is encoded in personal pronouns and definite articles before it is marked on indefinite articles (Mills 1985, Clahsen et al. 1994). Such findings support the assumption that children acquire individual form/meaning mappings item by item, or morphological class by morphological class (Tomasello 2003, Eisenbeiss 2003, Dressler 1997, 2005).

Some of the most detailed proposals regarding the time course of case acquisition come from research conducted in the framework of 'natural morphology'. Based on the salience and productivity of morphological forms and the mapping between these forms and their meanings in the input language, researchers determine which types of morphological forms children should prefer and acquire first. For instance, children's avoidance of zero suffixes, at a stage when they already have a good command of inflectional suffixes, may be explained by a preference for iconic mappings, where the morphological complexity of a linguistic form correlates with the complexity of its semantics. Therefore, zero suffixes in the genitive plural in Russian are anti-iconic and unnatural. Children tend to avoid such forms and replace them by non-existent but iconic forms (Ceytlin 1989).

The predictions of the natural morphology approach were tested in studies involving twenty morphologically rich and typologically different nominative–accusative languages, from the Slavic and Baltic, Finno-Ugric, and Germanic families. These studies suggest that the acquisition of grammatical categories involves three phases: premorphology, protomorphology and adult-like morphology (Dressler and Karpf 1995, Dressler 1997, Bittner et al. 2000, 2003, Voeikova and Dressler 2002, Dressler 2005: 9–10).

During the *premorphological* phase, children do not produce any morphologically marked forms, but use base forms or rote-learned forms. This happens frequently in children speaking agglutinating languages, e.g. early uses of fossilized

deictic particles like *siellä* 'there', täällä 'here', or local adverbs: *kotiin* 'home-ILL', *syliin* 'lap-ILL' in children acquiring Finnish (Laalo 2002: 93–4). This stage is comparable to the earlier stages without productive case-marking assumed in usage-based approaches or in generative structure-building approaches.

In many children, the precursors of morphological forms occur long before true case oppositions develop, e.g. a 'protomorpheme', or filler *-a* used as an inflectional ending that is a 'phonologically appropriate but semantically empty' element (Peters and Menn 1993: 743). Inappropriate *-a* marking was noted in the speech of several children acquiring Russian and Lithuanian (Savickienė 2002: 105, 113; 2003: 53; Gvozdev 1961: 164). Another process preceding the first grammatical oppositions is the use of (phonological) 'extra-grammatical operations', such as truncations, reduplications, blends, etc. (Kilani-Schoch et al. 1997: 15, 20–2). These processes are the first attempts to divide words into morpheme-like chunks.

First oppositions and 'miniparadigms' (oppositions of three and more forms of the same stem) signal the phase of *protomorphology*. Children start to detect the morphological patterns of the language being acquired. The first case oppositions emerge at different times even in children acquiring the same language. For example, in Russian the first case oppositions appear at the mean age of 31.5 months in boys and 27.8 months in girls according to more than 10,000 parental reports (Shapiro and Chistovich 2000: 52, 59). At this phase, the paradigms are still incomplete and many errors of omission and commission occur so that children's speech remains ungrammatical. This stage corresponds to the stage of lexically restricted generalizations that are assumed in usage-based approaches as well as in lexicalist versions of generative structure-building approaches (see section 24.4). Children start to use case forms in their prototypical semantic and syntactic functions, e.g. in Lithuanian the genitive is first used by children to mark possession and only later do other functions (such as the genitive of negation) develop (Savickienė 2002). Inflectional forms that are productive in child language are usually also those with higher type frequency in the input. Token frequency also plays a role, for example the personal name of a child is extremely frequent in most dialogues and its form may, thus, influence the individual process of the acquisition of case. Thus, for the Greek boy, Christos, his own name is the most frequent masculine noun he uses (Christofidou and Stephany 1997) and may help him learn the otherwise rare *-s* marker of masculine nouns.

At the phase of *adult-like morphology*, complete paradigms occur, but are initially restricted to frequent words. Children use case forms in several semantic functions but this stage is still not absolutely equivalent to the adult use of case since unproductive and rarely used forms receive erroneous case making. As proposed by Ceytlin, 'the system is learned before the norm' (Ceytlin 1989: 55–7).

Across languages, children acquire case at different points in time and at different rates. Case-marking in agglutinating languages occurs very early: the first case

978–0–19–920647–6 25-Spencer-ch24 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 382 of 383 May 16, 2008 18:59

oppositions in the speech of a Turkish child were observed at 1;3, and two Finnish children were found to utter their first case forms in appropriate contexts even earlier (at 1;0–1;1) (Laalo 2002: 89 ff.; Aksu-Koç and Ketrez 2007; Voeikova 2002). Children speaking Slavic and Baltic languages start with case oppositions at a relatively later age: the mean age of the first occurrences is about 1;9 (Savickienė 2002, 2003; Voeikova and Gagarina 2002). In languages with periphrastic case-marking on nouns and articles, the first contrastive forms emerge at the latest, between 2;2 and 2;3 (Klampfer and Korecky-Kröll 2002).

However, the chronological age at which first case forms emerge does not clearly predict future patterns of use (Voeikova 2002). An alternative measure, 'speed of development' (Emergence 2005) examines the monthly growth of paradigms in children's speech relative to the morphological richness of the input. Morphological richness is measured in the adult input by the mean size of the paradigms for nouns and verbs, and by the transparency, uniformity, and salience of word forms. When morphological richness is measured separately in the standard language and in children's input, it turns out that caregivers do not make use of the whole set of target forms. For instance, standard Russian possesses a richer set of nominal forms than German (Kempe and MacWhinney 1998), but caregivers typically use one or two oppositions of every given stem so that the difference in the mean size of nominal paradigm measured by the input is much more modest than the one presented in the adult grammar. Thus, the input that children get is not as typologically loaded as one may assume from the analysis of standard grammatical systems and may account for the time it takes for children to achieve adult-like productivity in case acquisition.

Taken together, cross-linguistic studies on the time course of case acquisition show a gradual development with dissociations between different cases and case markers. These observations are captured by approaches that assume incremental acquisition mechanisms.

24.6 Conclusions

As our discussion of theoretical approaches and empirical findings has shown, generative, natural morphology, cognitive-functional, and usage-based approaches may highlight different aspects of case acquisition and make different assumptions with respect to the exact role of nature and nurture in the acquisition process. However, all current models have to capture the observation that children produce systematic case errors stemming from syntactic, morphological, and semantic

978–0–19–920647–6 25-Spencer-ch24 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 383 of 383 May 16, 2008 18:59

factors. Purely semantic and purely syntactically based approaches to case acquisition offer interesting insights but cannot capture the cross-linguistic variability of case systems and acquisition paths or the incremental nature of children's case development. Therefore, many current approaches adopt probabilistic mappings between semantic and grammatical representations and consider the role of multiple factors (e.g. morphophonological properties of the input) that can influence the onset and speed of case development.

CHAPTER 25

CASE IN LANGUAGE PRODUCTION

ALISSA MELINGER
THOMAS PECHMANN
SANDRA PAPPERT

25.1 Introduction

Speech production involves the transformation of a to-be-expressed idea, or message, into lexical and grammatical content. This transformation involves the retrieval of the lexical items that correspond to the concepts comprising the intended message and the ordering of these lexical items into grammatical strings. In addition to stringing together content words, producing a grammatical sentence involves inserting the necessary function words and markers required to express grammatical aspects of the message such as tense (marked on the verb), definiteness (marked on a determiner), and grammatical function (marked on determiners, nouns, or in word order). Included in these steps is a process by which each argument of the verb is assigned a functional role within the sentence. This function can be indicated via case marking on the noun phrase. For example, when producing an active sentence in German, a retrieved argument that is assigned the functional role of direct object will be marked with accusative case. However, when producing a passive sentence, that same retrieved argument will be assigned the functional role of subject and marked with nominative case.

978–0–19–920647–6 26-Spencer-ch25 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 385 of 401 May 16, 2008 18:59

The question of how case assignment is achieved during sentence production has focused on two separate processes: case assignment and case realization. The latter is an element of morphophonological processing, which is achieved during the phonological encoding stage of utterance planning. In this chapter, we will focus on case assignment, which is achieved during the grammatical encoding stage of utterance planning. Questions about how case assignment is accomplished have been intimately intertwined with other questions, such as how the correct word order is achieved and how thematic roles are mapped to grammatical functions. Early proposals for sentence production models were highly influenced by the distribution and characteristics of naturally occurring speech errors. More recent revisions of these models have been further influenced by experimental investigations into structural and word order alternations using a method called **syntactic priming** (Bock 1989). In this chapter, we first lay out in gross terms the general views of the stages necessary for sentence production. We then focus on the evidence that has supported the various aspects of the models and how they directly or indirectly inform us about the processes responsible for case assignment in sentence production.

25.2 Sentence production models

Most language production models are essentially concerned with speakers' generation of single word (predominantly nouns) and single phrase (determiner + noun, adjective + noun) utterances, mostly ignoring longer syntactic stretches of speech. Since case assignment is basically a syntactic phenomenon specifying the role of arguments in a sentence, the processes underlying case assignment have largely escaped serious consideration by most theoretical models. The limited attention paid to sentence level phenomena has a primarily methodological basis (Butterworth 1980). Unlike in sentence comprehension where the input highly constrains the processes, it is extremely difficult to set up experiments such that the production of longer stretches of speech is predictable. Speakers' unconstrained descriptions of scenes, even very simple ones, vary both in formulation as well as in the time course of speaking.

Thus, early models of sentence production were built upon the systematic analysis of naturally occurring speech errors, such as the exchange error in (1).

(1) A maniac for weekends (intended utterance: A weekend for maniacs)

In her analysis of these errors, Fromkin (1971) noted that the inflectional elements associated with the exchanged nouns and the overall prosodic structure of the

phrase or clause remains intact. The observation that words, inflections, phonemes and prosodic structures can independently be involved in errors led to the development of the **Utterance Generation model** (Fromkin 1971; 1973), which proposes that sentence production unfolds over several steps each of which deals with different types of representations. According to the Utterance Generation model, the semantic message, namely the idea the speaker wishes to convey, gives rise to the generation of an abstract syntactic outline, which then guides lexical retrieval processes. Grammatical functions and case features are associated to positions within the syntactic outline. Thus, a word is assigned case when it is inserted into a particular position within the awaiting syntactic frame.

Garrett (1975; 1980) further developed the idea that utterance planning unfolds over several steps by drawing the distinction between **Functional** and **Positional** processes, which together comprise the larger stage of processing **Grammatical Encoding**. Most models now adopt some variant of this distinction (Dell 1986; Kempen and Hoenkamp 1987; Bock and Levelt 1994), although the specific make-up of each stage differs. In Garrett's original formulation, functional processing involves selecting the open class elements that correspond to the content elements of the message and assigning each retrieved argument a grammatical function within the sentence. Grammatical function assignment in Garrett's model is achieved by inserting the retrieved content elements into abstract syntactic slots, which are generated on the basis of the semantic message. To illustrate how such a model works, consider the possible syntactic alternatives to express a transfer event with three explicit arguments, e.g., GIVE (X, Y, Z). Such a message could generate several different abstract syntactic frames including the dative frame N_{nom} V N_{dat} N_{acc}. Retrieved nouns can only be inserted into the noun slots and retrieved verbs can only be inserted into verb slots. Lexical elements inherit the case features associated with the slots to which they are inserted.

The abstract syntactic frame indicates the grammatical functions of each slot within a clause but it does not specify order or hierarchical relations. These are specified during positional processing, which subsumes two subprocesses. The first subprocess is constituent assembly, or the construction of hierarchical constituent structures. The structures fix the order of the phrases as well as the dependencies within the hierarchy. These structures include part-of-speech labelled slots where the phonological forms of the content elements can be inserted. Positional processing is particularly important for languages with word order flexibility. It is responsible for producing topicalization, as in (2) or marked OVS word order, as in the German example in (3). The second subprocess subsumed under positional processing is inflection, which inserts function words and inflections into their appropriate positions within the hierarchy. Case marking affixes are phonologically encoded at this stage, commensurate with the abstract features passed down from the functional level.

(2) Him, I like.

(3) *Den Hasen frisst der Fuchs.*
The.ACC hare.ACC eats the.NOM fox

Another primarily syntactically driven model is Dell's **Spreading-Activation theory** of lexical retrieval (1986). Dell's model distinguishes three primary stages of processing: the syntactic, morphological, and phonological levels. The phonological encoding component of this model was extensively developed to account for phonological errors, but the same principles and mechanisms are proposed for syntactic and morphological processes as well. Each stage builds structures based on generative rules of the language and retrieves the relevant representations from the lexicon. The model uses a slot-and-filler mechanism to associate retrieved lexical representations with appropriate slots.

At the syntactic level, generative rules create syntactic frames, comprising an ordered sequence of category-labelled slots. The categories relate to the derivations of generative syntax, resulting in phrasal units, e.g. S, NP, VP, and terminal slots, e.g. N, V. Terminal slots are filled by the most active lexical item of the appropriate category. Dell's model makes little explicit mention of case assignment, but case features are presumably associated with structural positions within the generated syntactic frames. As in Garrett's model, functional role assignment is achieved by associating a lexical item with a structural position. However, in contrast to Garrett's model, the generated syntactic frames at the syntactic level already reflect the surface order of the intended utterance.

Although Dell's model is syntactically driven in that lexical selection is guided by the grammatical constraints of the to-be-filled structural position, the use of a particular structure building rule is sensitive to activation levels within the lexicon (Dell 1986: 316), thus demonstrating some elements of lexical guidance. Although Dell's model includes an explicit level at which stems and affixes are retrieved, he places inflection within the syntactic level. Thus, in a case marking language, a syntactic position for case-markers would be included within the syntactic frame. The phonological realization of the inflection, however, would be realized at the level of phonological encoding. For the correct case-marker to be inserted, case features have to be incorporated at the syntactic level guided by the specifications in the conceptual/message level.

Growing out of Dell's Spreading-Activation theory, Chang and colleagues (Chang et al. 2006; Chang 2002; Chang et al. 2000; Dell et al. 1999) developed several versions of their sentence production model, the most recent version of which is called the **Dual-Path Network**. The model was developed in part to account for the long-term component of syntactic priming (Bock and Griffin 2000) by linking adult syntactic behaviour to language acquisition mechanisms. Syntactic priming is the observation that speakers tend to recycle abstract syntactic structures that they

978–0–19–920647–6 26-Spencer-ch25 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 388 of 401 May 16, 2008 18:59

have recently processed. Bock and Griffin demonstrated that this tendency persists over as many as ten intervening utterances.

In the Dual-Path Network grammatical function assignment, inflection, and word order are essentially achieved together, in a single step. The message representation and compatible syntactic sequences conspire to retrieve lexical items in the order in which they should be mentioned. For example, a message consisting of an event with two arguments will initially make available both an active transitive and a passive structure. Since both structures begin with a noun phrase, the system will first try to retrieve one of the two arguments from the lexicon. Which argument is retrieved first will be largely determined by conceptual factors, such as saliency, prior discourse mention, animacy, etc. If the concept corresponding to the patient role is lexicalized first, then the passive structure will increase in accessibility relative to the active structure. Case marking of noun phrases is determined in tandem by the message system and the sequencer. Patients occurring postverbally in German will be preferentially marked with accusative case, as in (4), while patients occurring sentence initially are more likely to be marked with nominative case, since a common compatible structure is the passive structure, as in (5). The sequencer learns the appropriate case markings in the same way that it learns the grammatical strings of the language, namely during a training phase (cf. Chang et al. 2006).

(4) *Der* *Mann* *isst* *den* *Fisch.*
The.NOM man eats the.ACC fish

(5) *Der* *Fisch* *wird* *vom* *Mann* *gegessen.*
The-NOM fish is by. the-DAT man eaten

As in Garrett's (1975; 1980) and Dell's (1986) models, lexical retrieval is (partly) driven by syntactic options. Working in tandem with the content of the message, the sequencer determines which component of the message should be retrieved from the lexicon next. For example, given the restrictions of English syntax, the sequencer may decide to retrieve the main verb after having retrieved a word to express the agent of the event, since this is the most frequent sequence in English. The choice between syntactic alternatives is not sensitive to the lexical semantics or syntactic restrictions of specific lexical items. In fact, syntactic regularities are represented within the sequencing component (the syntactic parser), not within the lexicon.

In all of the models described above, an abstract syntactic frame constrains lexical retrieval processes; lexical items of the appropriate syntactic category are selected and inserted into the appropriate slots. However, it has been suggested that starting with a commitment to an abstract syntactic structure may not be sufficiently incremental. Specifically, an early commitment to a *specific* structure may not allow sufficient flexibility for the demands of the fast and fluent sentence production system. The issue of incremental processing was explicitly taken up in the **Incremental Procedural Grammar** of Kempen and Hoenkamp (1987). This

model, which was largely adopted by Levelt (1989), also incorporates two stages of sentence production, although the distribution of labour is somewhat different from its predecessors. The primary difference, however, is that the construction of syntactic structure is lexically driven, rather than syntactically driven. Like the original Garrett (1975) proposal, the first stage, the Lexico-Syntactic Stage, assigns hierarchical and functional relations between retrieved constituents. However, it differs from the original proposal and from the Bock and Levelt (1994, see below) proposal in that function words are inserted and word order is already computed in the first stage of processing. The second stage of sentence production, called the Morpho-Phonological Stage, is primarily concerned with linking phonological forms to sentence slots, calculating inflections, including case marking, and abstract intonational patterns. As mentioned above, the model also deviates from prior models in that it more strongly emphasizes the role of lexical items in syntactic planning.

The first stage of processing is the most relevant for purposes of case assignment. Kempen and Hoenkamp (1987) distinguish a number of different syntactic procedures. Categorical procedures, which can be triggered by the grammatical content of retrieved lexical items, build syntactic structures (e.g. noun phrases or verb phrases) and inspect the message for relevant feature specifications such as number, definiteness, or qualifications that require, for example, adjectives. Functional procedures handle the specifications required by these additional message components as well as handling the important task of determining the grammatical relations (e.g. subject, direct object, indirect object) between structures. The lexical entry of the verb is also the place where instances of idiosyncratic case assignment are specified.

To the extent that the syntactic rules of the language allow it, word order decisions will correspond to the order in which lexical items become available. Thus, the first noun to be retrieved will be placed in the left-most structural position. In this way, the model is both lexically driven, in that structure is built based on the grammatical requirements of the lexical items, and incremental, in that the order in which lexical items are retrieved influences sentence word order.

The model developed by Bock and Levelt (1994; Bock 1995) also adopts the differentiation between functional and positional processing within a lexically driven, incremental system. The subprocess most relevant to case production is again function assignment.

Within this model, function assignment is the process by which retrieved content words are assigned a grammatical function within the sentence. Furthermore, function assignment is claimed to be sensitive to thematic and information-structural roles in the preverbal message (Bock and Levelt 1994). It is also sensitive to the consideration of lexico-syntactic features (Bock and Levelt 1994; Bock 1995). In the network model of lexical access (cf. Levelt et al. 1999; Roelofs 1992, 2000), the syntactic (lemma) entry of a verb plays a central role. It specifies the subcategorization frame that maps the grammatical functions to configurational positions (Bock and

Levelt 1994). Once a syntactic function is assigned, it cannot be changed by syntactic processes like transformations. Consequently, the direct object of an active sentence does not correspond to the subject of a passive sentence with parallel meaning (Bock and Levelt 1994).

The model is also lexically driven in that constituent structure is created for lemmas as they are passed from the functional stage to the positional stage. In contrast to earlier two-staged models, Bock and Levelt (1994) do not posit abstract syntactic frames into which lexical items are inserted (e.g., Garrett 1975; Dell 1986). Rather, fragments of syntactic structure are built based on the demands of the retrieved lexical items and on the demands of the relations established during function assignment.

A model that does not subscribe to Garrett's distinction between functional and positional level processing is the **Tree-Adjoining Grammar**-based model of language production (Ferreira 2000; Ferreira and Swets 2002). This model conflates constituent structure assembly, word order, and functional role assignment into a single processing step. The model adopts a version of a Tree-Adjoining Grammar (Frank 1992; Joshi 1985; Joshi et al. 1975; Kroch and Joshi 1985) in which retrieved lemmas project their elementary trees, or the fragment of syntactic structure licensed by the particular lemma. Since determiners and nouns can only project as far as a DP, articulation cannot begin until the DP is assigned a grammatical role, which can only occur after the clause is projected by the retrieved verb. Thus, this model opts for lexically driven syntax over radical incrementality.

In the review of the above models, we have highlighted a number of ways in which the models differ from one another and also how the models' assumptions relate to grammatical function and case assignment. The models all share a rejection of structural transformations (however, see Franck et al. 2006 for evidence of movement operations) and all acknowledge that utterance encoding unfolds over several stages. All of the models also have difficulty accounting for non-canonical case assignments. For example, German has a set of ditransitive verbs that take two accusative arguments (6) despite the fact that the recipient argument behaves syntactically like an indirect object. Examples such as (6) demonstrate that case assignment cannot be completely derived from function assignment.

(6) … *dass Hans ihn etwas fragte*
… that Hans.NOM him.ACC something.ACC asked
'… that Hans asked him something'

The models differ, however, in a) how they distribute the labour over the various processing stages, b) the importance of incrementality, and c) whether functional roles are assigned and constituent structure is built in response to the requirements of retrieved lexical items or whether an abstract syntactic outline guides the lexical retrieval process. In the subsequent sections, we examine the evidence for each of these competing positions.

978–0–19–920647–6 26-Spencer-ch25 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 391 of 401 May 16, 2008 18:59

25.3 Evidence for the separation between functional and positional processes

Given the generally recognized separation of functional and positional processes, it has been argued that case assignment is within the domain of functional processes (or within Dell's syntactic stage). However, what is the evidence for such a separation and is there support for placing case assignment within the domain of functional processing?

One source of evidence for the separation comes from speech errors – their distribution, characteristics and scope, specifically phrasal exchanges. Speech errors have been systematically studied for decades and served as the basis for early models of sentence and single word production. We know that errors can occur at many levels of processing, from whole phrases to single phonological features. Early observations revealed that the pattern of phonological errors compared to word and phrasal errors differed in systematic ways. For one, exchanged segments usually originate within the same phrase while exchanged words usually originate in separate phrases. Second, word exchanges or substitutions usually respect part of speech constraints; nouns exchange with other nouns, not with verbs or adjectives. Likewise, nouns tend to exchange with or be replaced by other nouns of the same grammatical gender in gender marking languages (Marx 1999). When segments exchange, no such grammatical constraint is respected. These patterns led Garrett (1975, 1980) to propose the distinction between functional and positional processing, which was incorporated into many subsequent models.

Errors involving exchanges of phrases, rather than words, provide insights into functional assignment and case assignment. In an error such as (7) the exchanged pronouns display the correct case for their position within the sentence.

(7) you must be too tight for them (intended as They must be too tight for you, Stemberger 1982)

(8) Them must be too tight for you_{nom}

This is the general pattern observed for English (Stemberger 1982) and German (Berg 1987) pronoun errors. Bock and Levelt (1994) take this pattern as evidence that errors involving pronouns or other whole phrases are errors of functional assignment (and consequent case assignment), not of misordering the words during positional assignment. Thus, within the Bock and Levelt model, the second person singular pronoun was assigned the incorrect grammatical function (and case) during functional processing and then mapped into the structural position appropriate for that function. Within Garrett's model, the second person singular pronoun was inserted into the wrong slot within the abstract syntactic frame and

thus inherited from that slot the wrong case features. If the error had involved correct functional assignment but subsequent misordering, the error would be as in (8). Some errors of the latter type in (8) have been reported for German. Levelt (1989) discusses the example in (9) from the Bierwisch corpus. This unusual positional level error was observed rather than the more typical example involving functional role assignment, which would give rise to the error in (10). The contrast between the unusual error in (9), in which case occurs with the correct word but in the wrong position, and the more common error in (10), in which case occurs with the wrong word in the correct structural position, highlights the need for the functional–positional level contrast and for locating case assignment in the former. Note also that these errors are very different from the type observed by children (see Eisenbeiss, Narasimhan, and Voeikova, this volume) and by aphasic speakers (see Lamers and Ruigendijk, this volume).

(9) *Wenn der Wand aus der Nagel fällt* [intended as *Wenn der Nagel aus der Wand fällt*]
When the.DAT wall from the.NOM nail falls (intended as When the.NOM nail from the.DAT wall falls)
'When the wall falls out of the nail . . .' (intended as 'When the nail falls out of the wall . . .')

(10) *Wenn die Wand aus dem Nagel fällt.*
When the.NOM wall from the.DAT nail falls . . .
'When the wall falls out of the nail . . .'

(11) He wants us to do something else (intended as We want him to do something else; discussed in Bock 1995: 187)

A second source of evidence that phrasal errors occur at the level of functional assignment is that verbs agree with their actual subjects, not their intended subjects, as in (11). This suggests that function assignment, the determination of which element is subject of the sentence and thus controls agreement processes, was incorrectly implemented. If function assignment had been correctly implemented but the constituent structure incorrectly assembled, then we would expect the verb to agree with the wrong element in the actual utterance.

Experimental evidence for the separation between functional and positional processing stages also comes from **syntactic priming** studies investigating speakers' preferences for one of two alternates of a diathesis alternation. These studies take advantage of the observation that speakers tend to reuse previously encountered structures (e.g. Kempen 1977; Tannen 1989). Several variants of the method exist, varying from sentence recall (Potter and Lombardi 1990; Ferreira and Dell 2000) and sentence completion (Pickering and Branigan 1998; Corley and Scheepers 2002; Hartsuiker et al. 1999b) to extemporaneous picture description (Bock 1986, 1987; Bock and Griffin 2000; Branigan et al. 2000). In the commonly used picture

description variant, sentence primes precede target pictures, which speakers describe. Prime sentences instantiate one of the possible structural alternatives for the description, e.g., either what is typically called a Double Object construction, as in (12), or a Prepositional Object construction, as in (13).

(12) The policeman issued the driver a ticket.

(13) The policeman issued a ticket to the driver.

Target pictures are generally semantically unrelated to the prime sentence but can be described using either the primed construction or an alternative construction. The findings from these studies demonstrate that speakers reuse the structure presented in the prime sentence (see Pickering and Branigan 1999, for an early review of the results obtained with this and related paradigms). As will be clear from the examples below, one early goal of syntactic priming studies was to demonstrate that syntax had representations that could be distinguished from specific lexical items or specific meanings.

Depending on the nature of the alternation and the relationship between the prime and target sentences, the results can be attributed to either functional or positional levels. For example, Bock and Loebell (1990) found that active primes with a locative phrase, as in (14), induced speakers to produce passive rather than active picture descriptions. This was a surprising result because active locatives and passives do not share a mapping of thematic roles to grammatical functions. Thus, the priming effect cannot be due to a common process of function assignment. However, the two sentence types do share a common constituent structure and therefore Bock and Loebell attributed the priming effect to positional level processes, where constituent structure is assembled.

(14) The plane was landing by the control tower.

Furthermore, Hartsuiker et al. (1999b; Hartsuiker and Westenberg 2000; Smith and Wheeldon 1999; Wheeldon and Smith 2003) demonstrated that alternations in word order can also be primed, even when they do not involve contrasts in functional assignment or constituent structure. They found word-order priming when comparing the alternation between figure–ground and ground–figure orders, such as (15) compared to (16). These results were interpreted as strong evidence for priming effects at the positional level, since these sentences do not differ in functional assignment. Moreover, data from Japanese add to the evidence of positional priming. Japanese displays free word order. Dative sentences do not alternate between Double Object and Prepositional Object constructions but syntactic functions are signalled by case marking. In a primed production experiment, effects of linear order were found that are not attributable to processes at the functional level (Yamashita and Chang 2006). Specifically, target utterances such as (19) were more common following prime sentences like (17), which have the same order of

978–0–19–920647–6 26-Spencer-ch25 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 394 of 401 May 16, 2008 18:59

arguments, and less common following prime sentences like (18), which have the reversed order of dative and accusative arguments.

(15) *Een boek ligt op de plank.*
'A book lies on the shelf.'

(16) *Op de plank ligt een boek.*
'On the shelf lies a book.'

(17) *Akiko-wa kagi-o tomodati-ni ageta.* (Prime Sentence 1)
Akiko-TOP key-ACC friend-DAT gave
'Akiko gave the key to the friend.'

(18) *Akiko-wa tomodati-ni kagi-o ageta.* (Prime sentence 2)
Akiko-TOP friend-DAT key-ACC gave
'Akiko gave the friend the key.'

(19) *Taro-wa hana-o onnanoko-ni okutta.* (Target sentence)
Taro-TOP flowers-ACC girl-DAT sent
'Taro sent flowers to the girl.'

While the examples from Dutch and Japanese demonstrate that structural priming does not need to involve contrasting function assignments, priming of either the double object or prepositional object structures in languages like English must be attributed to functional level processes, as function and case roles distinguish (and determine) the two structures. Thus, structural priming provides evidence that different types of processes can be primed and produce similar results, namely the persistence of a surface structure. Thus, these studies demonstrate that structural priming can influence a) word order alternations, b) the building of constituent structure, and c) the assignment of functional roles.

Structural priming can clearly target functional level processes, but is it sensitive to case assignment? Melinger and Cleland (2005; in preparation) investigated the influence of case assignment on the priming of noun phrase structure. Cleland and Pickering (2003) showed that the preference for adjectival noun modification (*the red cow*) vs. relative clause modification (*the cow that is red*) was also sensitive to persistence effects. Melinger and Cleland followed-up on this observation; they investigated whether the role that the noun phrase plays within a sentence contributes to this persistence effect. In their study, German native speakers described the spatial location of a figure with respect to a ground object, as in (20). Sentences included either a modified figure (21) or ground (22).

(20) *Der laufende Clown / Clown, der läuft, ist unter dem Tisch.*
The.NOM running.NOM clown / clown who.NOM runs is under the.DAT table
'The running clown / clown who is running, is under the table.'

(21) *Der stehende Arzt ist links von der Nonne.*
The.NOM standing.NOM doctor is left of the.DAT nun.
'The standing doctor is left of the nun.'

(22) *Links von dem stehenden Arzt ist die Nonne.*
Left of the.DAT standing.DAT doctor is the.NOM nun
'Left of the standing doctor is the nun.'

In German, these roles are explicitly case-marked on the determiner and can also be marked on the adjective and the noun. Results demonstrated a robust structural priming effect that was insensitive to the case of the modified NP: descriptions included more relative clause modifications following prime sentences with relative clause modifications, irrespective of the role that the modified noun played in the prime sentence. Thus, a nominative case-marked figure NP did not prime another nominative figure NP more than a dative ground NP. This finding is consistent with early results demonstrating that priming is unaffected by thematic role mismatches between prime and target sentences (Bock 1989; Bock and Loebell 1990). It also could support a general (context-free) process for constructing noun phrases which is insensitive to the larger context in which the noun phrase is embedded (cf. Branigan et al. 2006; Apel and Melinger 2005, in preparation). What seems clear is that sharing case features does not enhance the NP priming effect.

To summarize, the speech error data seem to clearly require a level where grammatical relations are established (and where case is assigned) and where phonological elements, stems and inflections, are inserted into hierarchically organized frames. These patterns served as the foundation for the two-staged models. It is also clear that persistence effects arise at any and all of the various stages of processing. However, direct evidence from structural priming that distinct stages of processing are required is sparse (for evidence against the distinction, see Pickering et al. 2002). Nevertheless, the findings are consistent with a functional level locus of case assignment.

25.4 Evidence for lexical guidance (or verb primacy) in functional assignment

Above we have examined the evidence for two-staged models of sentence production and located case assignment within the earlier stage, namely functional processing. We now turn our attention to the principles that drive the assignment

of case. Specifically, we present evidence for whether function and case assignment are primarily message- (or syntax)-driven or lexically driven.

Within the Bock and Levelt (1994) model of sentence production, functional assignment, and therefore case assignment, is strongly influenced by the lexical-specific features of the sentence head. Yet, at the same time all models aspire to capture the incremental nature of speaking. Thus the tension between lexically driven syntax and incremental production has been growing for the last few decades. We first focus on the evidence for lexical guidance, the influence of lexical features on the unfolding syntax, and then turn, in the next section, to the evidence for incrementality.

First, the terms lexically driven and lexical guidance have been used in various ways within the production literature. Our intention is to contrast lexically driven with syntactically driven. In lexically driven sentence production, syntactic structures are built based on the grammatical information associated with retrieved lemmas. Thus, if a noun is retrieved, an NP is constructed to house it. If an intransitive verb is retrieved, then a VP with no internal arguments is constructed. In syntactically driven sentence production, the roles are reversed. A noun is retrieved because the next free position in the sentence under construction requires a noun. An extreme reading of the lexically driven position highlights the primacy of the verb in licensing clauses. Such a reading suggests that, until a verb is retrieved to guide function assignment, positional processing cannot begin. An alternative use of the term lexically driven relates to incrementality. Specifically, sentence production can be said to be lexically driven if the order of mention of elements in a sentence is determined by the order in which concepts become lexically available. This notion will be picked up again in the section on incrementality.

Evidence for lexically driven sentence production and the centrality of the verb in functional processing comes partly from linguistic observations that, for example, case can be assigned idiosyncratically by different verbs and the pattern cannot be easily attributed to semantics. Likewise, semantically similar verbs can accompany different sets of structural options. However, investigations into the relationship between conceptual features like argument structure and grammatical features like case underplay the central role of lexical features. For example, in German, overt case marking paired with a sentence-final verb position provides the opportunity to investigate argument structure preferences and their impact on case assignment. Scheepers and Corley (2000) report a sentence completion study in which they presented sentence fragments beginning with a subject and an object (both referring to an animate entity; Scheepers, p.c.). Case of the object was varied. After an accusative object, there was a prevalence of single object completions whereas after a dative object, more double object completions were produced. This effect might be reduced to simple word order preference: Datives tend to precede accusatives. However, Pappert et al. (2007) conducted a corpus and a completion study focusing on the impact of case and animacy on the production of such sentences. They

found more single than double object sentences in all case and animacy conditions (including that with an animate or inanimate nominative followed by an inanimate dative). They only found more double object than single object structures after a nominative and a dative NP, both referring to an animate entity. A closer inspection of the completion data revealed a characteristic distribution of the produced verb types: fragments including animate datives mostly elicited transfer verbs whereas elsewhere, verbs of causal affection and psych verbs predominated. Since the verbs in these studies were sentence final, the completion preferences must have been driven by conceptual features. Thus, the results are interpreted as evidence of a strong association between case, animacy, and thematic roles.

Experimental evidence for the influence of verb-specific information comes from syntactic priming studies. Priming effects are observed to be larger when the prime and the target share a common verb (Pickering and Branigan 1998), but not when they share other lexical elements such as prepositions (Bock 1989), or aspectual markers (Pickering and Branigan 1998). To explain the lexical boost, Pickering and Branigan argue that structural priming arises from the reselection of combinatory information, similar to verb subcategorization frames, associated with the verb's lemma representation. However, recent work from several independent laboratories suggests that the lexical boost and syntactic priming arise from two separate mechanisms (Konopka and Bock 2005; Melinger and Cleland, in prep). While the exact nature of the two mechanisms is still unclear, evidence for a distinction comes from the time course of the two effects. Specifically, syntactic priming, with no lexical overlap is very long lasting, with syntactic priming effects persisting over even ten intervening sentence trials (Bock and Griffin 2000). In contrast, the lexical boost seems to be more transient, diminishing after even one intervening trial (Konopka and Bock 2005; Melinger and Cleland, in prep). Other evidence for the lexical influence on syntax comes from investigations of verbs that do not participate in the relevant diathesis alternation. Specifically, Melinger and Dobel (2005) used a variant of the syntactic priming method described above to investigate lexical involvement in the dative alternation. In their experiments, sentence primes were replaced with single verb primes. The prime verbs were non-alternating dative verbs such as *ask*, which is only acceptable in the ditransitive structure, and *donate*, which is only acceptable in the double object construction. Melinger and Dobel observed that speakers described unrelated events with the structure compatible with the prime verb, although no sentence for the verb was presented or relevant to the study. Their results suggest that one source of syntactic priming is the selection of verb-specific subcategorization frames which in turn license the construction of a certain constituent structure.

If sentence formulation is strictly lexically guided, then perhaps the main verb is retrieved prior to its arguments, so that functional assignment and case assignment can proceed smoothly. There is some evidence that speech onset waits for the retrieval of the main verb (Kempen and Huijbers 1983; Lindsley 1975, 1976). Using

a picture description task, Kempen and Huijbers compared speech onset times for naming just the agent within a scene, just the action, the agent + action and the action + agent. Actions took longer to name than agents and naming times for both complex phrases were as long as the action naming condition. This suggests that speech onset is delayed until both the agent and the action have been retrieved. However, subsequent picture naming studies using distractor words that are semantically related to the sentence verb suggest a more moderate view. Namely, if the verb comes early in the sentence, it can drive functional assignment but if it comes later in the sentence, functional assignment will commence on the basis of conceptual features (Schriefers et al. 1998). Additional evidence for such an alternative route comes from the correlation between the semantic features of the verb and its arguments and the mapping of arguments to grammatical functions (Bock and Levelt 1994; Bock et al. 1992; Ferreira 1994). Despite the absence of semantic interference effects when the verb is not sentence initial, there is some evidence for advanced phonological planning of intransitive verbs. Schnur et al. (2006) asked speakers to produce simple intransitive sentences while ignoring distractor words that were phonologically related or unrelated to the main verb. Although the verb was not sentence initial, reaction times to begin uttering the sentence were faster when the distractor word was phonologically related to the verb compared to when it was unrelated, supporting advanced phonological planning of the verb. Unfortunately, given cascading activation from lemmas to wordforms (e.g. Peterson and Savoy 1998), advanced phonological planning is consistent with, but not unequivocal evidence for, advanced verb selection.

To summarize, there is evidence to support the claim that individual lexical items include grammatical specifications and that the parser is sensitive to these specifications, but there is little evidence to suggest that the parser waits for this information to become available before it starts to assign grammatical functions and build syntactic structure.

25.5 Evidence for and against (radical or weak) incrementality

Although issues of incrementality are not directly relevant to how and when case features are assigned, it is important with respect to the issues of lexical involvement in functional processing. Clearly, strict lexical guidance as we have defined it here is at odds with radical incrementality, which holds that conceptually available message components become lexicalized before less available components and the retrieved lemmas are automatically mapped to early sentence positions. In other words, the

first words to be retrieved will be the first uttered without any planning of the overall syntactic structure of the sentence (De Smedt 1990; Kempen and Hoenkamp 1987; Schriefers et al. 1998; van Nice and Dietrich 2003). Strict lexical guidance is also inconsistent with a weak incremental view (Ferreira and Dell 2000; Bock and Levelt 1994; Christianson and Ferreira 2005), which suggests that conceptual accessibility influences not just word order irrespective of the overall structure but also interacts with grammatical role assignment processes. In other words, the claim is not simply that conceptually accessible message components are lexicalized earlier, but that the availability of particular arguments triggers particular syntactic mechanisms which map the arguments to grammatical functions higher on the grammatical function hierarchy, which tend to occur earlier in the sentence.

The complication which arises from strict lexically driven syntax is that in many languages the verb comes very late in the sentence. If the verb must be selected prior to its arguments and prior to the construction of constituent structure, the verb must be held in a buffer while it waits for its turn to be uttered. This goes against the very grain of incremental production. It also suggests an early commitment to a verb and, potentially, to one of that verb's compatible subcategorization frames. In essence, lexically driven syntax predicts that alternative structures compatible with the selected verb will compete for selection. However, the incremental view predicts that more structural options allow structural decisions to be made on the fly. Evidence in support of both lexically driven syntax and incrementality has been observed.

Ferreira (1996) argued that resolving the competition between possible structures should be time consuming. Thus, sentences containing verbs with multiple compatible structural options, such as the alternating verb *give*, should require more time to be produced than similar sentences with non-alternating verbs, like *donate*. However, what he observed was that alternating verb sentences were produced faster than non-alternating verbs. The ordering flexibility afforded by the alternating verbs meant speakers could make word order decisions and structural adjustments on the fly, mentioning either the patient or the recipient first depending on whichever was more accessible. Non-alternating verbs did not allow such on-the-fly decision making. Thus, speakers had to commit earlier, resulting in longer speech onset latencies.

In contrast, Stallings et al. (1998) found evidence both for lexical guidance and competition between structural alternatives. Stalling et al. investigated heavy-NP shift (Kimball 1973) in English, namely the preference for longer constituents whose canonical position is adjacent to the verb to occur in clause-final position. They observed that the disposition of the verb to allow for non-adjacent direct objects strongly influenced the rate with which speakers shifted heavy NPs. Specifically, verbs which subcategorized for either a direct object or a sentential complement allowed for more shifted structures than verbs that only subcategorized for a direct object, since prepositional phrases often precede the sentential complement, see

(23). Competition was observed in the speech onset latencies. Before speaking, speakers indicated via button press which structure they intended to use (Experiment 1). Decision latencies reflected competition between available structures; when the verb's disposition supported both orders, decision times were longer than when the verb supported only one order.

(23) Mary said **in a loud voice** that Bill would sing.

Corpus studies provide additional evidence that the production system is finely tuned to the probabilistic preferences of verbs (e.g. Gries 2005; Jaeger and Levy 2006; also see Bresnan 2006 for evidence that speakers' grammaticality or acceptability judgments are sensitive to probabilistic forces imposed by the larger discourse). Gries (2005) demonstrated that structural repetition effects in corpora were verb-specific: some verbs with strong structural preferences were resistant to priming while others were more receptive. Experimental and corpus evidence consistent with an early commitment to a verb, and hence its selectional preferences, promotes a version of lexically driven syntax. Sentence production models that give precedence to verbs provide an easy solution to the problem of irregular case assigning verbs (e.g. quirky case).

25.6 Conclusions

In this chapter we presented a description of some models of sentence production and their theoretical treatment of function and case assignment. We contrasted more lexically driven two-staged models (Bock and Levelt 1994; Bock 1995; Kempen and Hoenkamp 1987) with models that are syntactically driven (Chang 2002; Chang et al. 2000, 2006; Dell 1986; Dell et al. 1999; Garrett 1975). We presented the evidence for distinguishing between Functional and Positional processes and for locating case assignment in the former. We highlighted how different views regarding the involvement of lexical representations in syntactic processes and the importance of incrementality influenced the developing treatment of case.

The models discussed were developed over a period of over thirty years and they reflect, to some extent, the state of the field at various points in time. The models were largely informed by investigations of Germanic languages, such as English and Dutch, which have rather impoverished case systems. Over the years, direct investigations of case assigning processes have been few and far between. Thus, our review focused greatly on related phenomena. Although no consensus has yet emerged, there is growing agreement that incrementality is vital for fast and fluent speech and perhaps the isomorphism between syntax and semantics

can be harnessed to explain many of the results previously attributed to lexical involvement in functional processes. Yet, despite this trend, the importance of verb-specific representations for syntactic processes still holds a dominant place in the literature. Future investigations that shed light on the production of verb final structures or on the influence of verb-specific idiosyncrasies will be crucial to our further understanding of these processes.

Finally, we should note that all of the reviewed models were developed primarily on the basis of English, which does not have a particularly rich case system, and somewhat influenced by German. Nothing at all is known about how other systems of case are achieved. However, recent studies using the syntactic priming method have started investigating underrepresented languages including Japanese and Korean. Extending the investigation to languages with more cases or different systems, e.g. ergative/absolutive systems, is crucial to the development of a more explicit model of case assignment within language production.

Acknowledgements

The authors would like to thank Franklin Chang, Merrill Garrett, Gerard Kempen, and the internal reviewer of the chapter for helpful comments and discussion. All errors and omissions are our own.

CHAPTER 26

CASE IN LANGUAGE COMPREHENSION

MARKUS BADER
MONIQUE LAMERS

26.1 Introduction

Research on human language comprehension has been heavily influenced by properties of the English language. Since case plays only a minor role in English, its role for language comprehension has only recently become a topic for extensive psycholinguistic investigations. In the present survey of this research, we will concentrate on the comprehension processes that are responsible for syntactic analysis. In the psycholinguistic literature, these processes are called the human parsing mechanism or the human sentence processing mechanism (HSPM). In most general terms, the task of the HSPM is to integrate each word of an input sentence into a complete syntactic structure. According to the Strong Competence Hypothesis (cf. Bresnan and Kaplan 1982), the syntactic structures computed by the HSPM are exactly those structures that are specified by the competence grammar. Given this hypothesis – which we will assume for the rest of this article – syntactic theory tells us what the HSPM has to compute. How the computations proceed is the topic of psycholinguistics.

Given the plentitude of approaches to syntax (as witnessed by the contributions to part I of this volume), adopting the Strong Competence Hypothesis is much less restraining than one might have hoped. Most of what we say in this review is

978–0–19–920647–6 27-Spencer-ch26 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 403 of 418 May 16, 2008 18:59

neutral with regard to particular syntactic assumptions. To hold the presentation as concrete as possible, we will assume that the HSPM computes phrase-structure representations enriched by various syntactic features, in particular case features on NPs.

We can now state the task of the HSPM more precisely. Following Mitchell (1994), we will distinguish between processes of structure assembly and processes of structure checking. The assembly processes are responsible for building up a phrase-structure representation by integrating each upcoming word into an unfolding phrase-structure tree. The purpose of the checking processes is 'to ensure that any structure under scrutiny satisfies all the detailed linguistic constraints that have to be met in well-formed and felicitous sentences' (Mitchell 1994: 401). They include the checking of subject–verb agreement as well as other agreement relations, the appropriate matching of arguments to argument structure positions, and the checking of case requirements imposed by verbs. Checking processes will thus be a major topic of the current survey. Checking processes have been investigated much less intensively than processes of structure assembly. This picture is changing however, with more and more languages becoming the topic of psycholinguistic investigations.

The organization of this chapter is as follows. The next section provides a short introduction into current research concerned with the HSPM. In section 26.3, we discuss how syntactic functions are assigned in the face of morphological case ambiguity. Section 26.4 takes a closer look at the role of case for identifying clause boundaries in languages like Japanese and Korean. The main focus of sections 26.3–4 will be the problem of syntactic ambiguity resolution. Sections 26.5 will discuss whether markedness distinctions that have been postulated to obtain between different cases are reflected in language comprehension. The final section 26.6 gives a summary of the chapter.

26.2 How case might help the Human Sentence Processing Mechanism (HSPM)

This section introduces some background information regarding the HSPM (for more comprehensive introductions, see Mitchell 1994; Pickering 1999; Tanenhaus and Trueswell 1995). As was said above, at least for the languages that will be considered in this article a major task of the HSPM consists in integrating each upcoming word into an unfolding phrase-structure representation, with integration

978–0–19–920647–6 27-Spencer-ch26 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 404 of 418 May 16, 2008 18:59

comprising both processes of structure assembly and various checking processes. An important obstacle to the integration task is the problem of ambiguity. Often a word can be integrated in more than one way, as in the well-known example in (1).

(1) a. In order to help *the little boy* put down the package he was carrying.
b. In order to help *the little boy* Jill put down the package she was carrying.

When processing sentences (1) from left to right, an ambiguity arises when the phrase *the little boy* is encountered. This phrase could be integrated either as the object of the preceding verb *help*, or as the subject of the upcoming main clause. This ambiguity is only a local one. In (1a), the immediately following main-clause verb *put* unambiguously requires the NP *the little boy* to be its subject. In (1b), in contrast, the following NP *Jill* makes it clear that the NP *the little boy* must be the object of *help*. Both intuitions and experimental results show that reading a sentence like (1a) causes severe processing difficulties (cf. Frazier and Rayner 1982; Sturt et al. 1999). When reading this sentence, there is a strong preference to analyse the ambiguous NP as object of the preceding verb. Since this is not compatible with the remainder of (1a), readers are garden-pathed. Sentence (1b), in contrast, continues in a way that is compatible with the initial preference for *the little boy*, and this sentence is accordingly processed smoothly.

There are several ways to describe the local ambiguity in (1). First, this is an attachment ambiguity. The NP *the little boy* can either be attached to the VP headed by *help*, or to the upcoming main clause. Second, this is a syntactic function ambiguity. The ambiguous NP is either an object or a subject. Third, the local ambiguity seen in (1) is an ambiguity as to the boundary between subordinate and main clause. The ambiguous NP is either part of the initial embedded clause or of the main clause. These descriptions are not independent. For questions of sentence comprehension, this has the important consequence that it is a priori not clear what kind of information is responsible for the strong preference visible in (1). For example, is the preference for attaching *the little boy* as an object of *help* a preference for a certain phrase-structural configuration, or is it a preference for a certain case assignment?

An answer to the question of whether ambiguities like those in (1) are resolved in terms of phrase-structure configurations or in terms of case has obvious consequences for the role unambiguous case morphology might have for sentence comprehension. For example, what happens when we replace the ambiguous NP *the little boy* by NPs which are unambiguously marked for case, as in (2)?

(2) a. In order to help he put down the package he was carrying.
b. In order to help him Jill put down the package she was carrying.

This question brings us to one of the most controversial issues with regard to syntactic ambiguity resolution: What types of information guide the initial decisions of the HSPM? With respect to case information, an experiment conducted by Traxler and Pickering (1996) is one of the few experiments that directly addressed this question using English sentence material (cf. Trueswell et al. 1993). Sample sentences from this experiment are provided in (3).

(3) a. I recognized (that) *you and your family* would be unhappy here.
b. I recognized (that) *she and her family* would be unhappy here.

The main findings of Traxler and Pickering (1996) can be summarized as follows. When no complementizer was present, readers spent significantly more time reading the unambiguously case-marked NP than the case-ambiguous NP; no significant difference showed up when the complementizer unambiguously signalled the beginning of the complement clause. Since Traxler and Pickering found this pattern of results already in reading time measures capturing early aspects of language comprehension, it can be concluded that case morphology affects the HSPM quite rapidly. However, as Traxler and Pickering discuss in detail, the results are compatible with either phrase-structure information being processed before case information, or vice versa.

According to the first possibility, the increased reading times for the case-unambiguous phrase *she and her family* would come about as follows. When the word *she* is recognized, initially only the lexical category of the word is seen by the HSPM. Since *she* is an NP and the simplest syntactic structure to build at this point is one where *she* is the object of the preceding verb, *she* will be attached as the object of *recognized*. Only after integration is completed will processes of case checking start their work. These processes will quickly discover that *she* has not the case required of an object, which will lead to a slow-down of the reading process because the initial decision about the syntactic analysis of *she* has to be revised.

Alternatively, one might assume that case information can govern the initial attachment decisions of the HSPM. Under this assumption, the HSPM would never falsely attach the NP *she and her family* as an object to the preceding verb. Instead, immediately on encountering *she*, the HSPM would know that this is a nominative-marked NP, and that therefore a new clause has to be opened in order to create an attachment position for a subject. Although the HSPM would not make any temporary mistake under this account, the increased reading times on *she and her family* in comparison to *you and your family* could still be explained easily. After all, integrating an object immediately after *recognized* involves less structure building than integrating a subject because only in the latter case will it be necessary to postulate a new clause.

To sum up thus far, the experiment by Traxler and Pickering (1996) shows clear evidence that case information is used at an early point by the HSPM, but at

978–0–19–920647–6 27-Spencer-ch26 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 406 of 418 May 16, 2008 18:59

the same time it also reveals a major difficulty in determining the exact timing relation between different types of linguistic information. According to the first account of their results – the garden-path account – phrase-structural information has priority over case information, whereas the reverse is true under the second, the complexity account. As we will see later, such problems of data interpretation occur even for languages in which case plays a much more prominent role than in English.

A second point which emerges from the preceding discussion is that there are at least two ways in which case morphology might help the human parsing mechanism: it might help in identifying syntactic functions, and it might help in finding clause boundaries. These two possibilities were confounded in the examples considered so far, but this was mainly due to the rather rigid word order of English which has the effect that ambiguities with respect to the syntactic function of a phrase are rarely encountered within a single clause (there are some systematic exceptions to this claim which we will consider in due course).

As soon as we broaden our language base, we can see that the two possible functions of case – identifying grammatical roles and identifying clause boundaries – can exist independently. The problem of identifying syntactic functions has been extensively investigated for Dutch and German where it is much more widespread than in English. A typical example from German is provided in (4).

(4) a. *Peter hat Maria gefallen.* Either SO or OS
P. has M. pleased
'Maria pleased Peter.' or 'Peter pleased Maria.'

b. ***Der*** *Peter hat* ***der*** *Maria gefallen.*
the.NOM P. hat the.DAT M. pleased
'Peter pleased Maria.'

c. ***Dem*** *Peter hat die Maria gefallen.*
the.DAT P. hat the.NOM M. pleased
'Maria pleased Peter.'

Sentence (4a) is a globally ambiguous sentence, in which either the first NP is the subject and the second the object, or vice versa. The two readings of (4a) can be unambiguously signalled if its two proper names are used together with a determiner, which is grammatically licit in German. (4b) gives the corresponding subject-before-object (SO) clause, (4c) the corresponding object-before-subject (OS) clause. We will discuss the major findings on syntactic function ambiguities in section 26.3. Afterwards, we will discuss the problem of identifying clause boundaries as it is found in languages like Japanese and Korean.

978–0–19–920647–6 27-Spencer-ch26 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 407 of 418 May 16, 2008 18:59

26.3 Processing Syntactic Function Ambiguities

When discussing English above, we said that there are some systematic exceptions to the claim that ambiguities concerning the syntactic function of phrases are normally not found within single clauses. Here is one of these exceptions (from Stowe 1986).

(5) a. Who = Object of verb
My brother wanted to know **who**$_i$ Ruth will bring ___ home to Mom at Christmas.
b. Who = Object of preposition.
My brother wanted to know **who**$_i$ Ruth will bring us home to ___ at Christmas.

The two sentences in (5) contain the dislocated wh-phrase *who*. The canonical position of the phrase has been marked by '___'. Following Fodor (1978), it has become customary within psycholinguistics to call a displaced phrase **filler** and the canonical position with which the filler has to be associated **gap**. The task of the HSPM is to correctly identify the gap corresponding to the filler. As shown by (5), the location of the gap can be locally ambiguous. The most important finding of Stowe (1986) was that a garden-path effect occurs at the position of *us* in (5b) (cf. Crain and Fodor 1985). It seems that the HSPM is expecting the gap for *who* directly after the verb *bring*, and is therefore surprised if the position after *bring* is already filled by another phrase. The resulting garden-path effect has accordingly become known as the filled-gap effect.

Ever since Fodor (1978) introduced the notion of filler–gap processing into psycholinguistics, it has been a topic of active research. In one of the studies that were seminal in bringing a cross-linguistic perspective into psycholinguistic research concerned with the HSPM, Frazier (1987) started the investigation of subject–object ambiguities in Dutch under the topic of filler–gap processing (under a different perspective, this kind of ambiguity was already part of an experiment by Nooteboom et al. 1978). An example sentence from Frazier (1987) is given in (6).

(6) *Jan houdt niet van de Amerikaanse die de Nederlander wil*
Jan likes not of the American who the Dutchperson wants
uitnodigen.
invite
'Jan does not like the American . . .
. . . who wants to invite the Dutchperson'
. . . who the Dutchperson wants to invite'

Sentence (6) is globally ambiguous. As indicated in (7), the ambiguity revolves around the location of the trace associated with the relative pronoun *die* ('who'). Either the trace precedes the second NP *de Nederlander* ('the Dutchperson') and the relative clause has a SO structure, or the trace follows the second NP, which gives rise to an OS structure.

(7) a. $[_{CP}\ [_{Rel-Pronoun}\ \text{die}]_i\ [_{IP}\ t_i\ [_{NP}\ \text{de Nederlander}]\ \ldots$
b. $[_{CP}\ [_{Rel-Pronoun}\ \text{die}]_i\ [_{IP}\ [_{NP}\ \text{de Nederlander}]\ t_i\ \ldots$

In an experiment where participants had to indicate which of the two possible structures they had preferentially understood when reading globally ambiguous sentences like those in (6), Frazier (1987) found a preference for the SO-structure. That is, the structure with the trace in initial position was the structure preferred by participants. This finding led Frazier (1987) to postulate the **Active Filler Hypothesis** (AFH) which we give below in the formulation of Clifton and Frazier (1989).[1]

(8) Active Filler Hypothesis (AFH)
When a filler of category XP has been identified in a non-argument position, such as COMP, rank the option of assigning its corresponding gap to the sentence over the option of identifying a lexical phrase of category XP.

Note that the AFH also explains the filled-gap effect illustrated in (5). When encountering *who*, the HSPM identifies this phrase as an active filler. The preferred position of its gap is the object position directly after the verb according to the AFH. If this position turns out to be occupied by another phrase, as in (5b), the filled-gap effect occurs. The AFH is thus a prime example of a phrase-structure parsing principle which does not only explain phrase-structure parsing as such but also certain facts about the preferred assignment of case features. Importantly, however, the AFH does not mention case at all. It is a mere side effect of the AFH that in the Dutch example in (6), the relative pronoun is preferentially assigned nominative case and the following NP accusative case.

Frazier's (1987) study initiated a comprehensive research effort devoted to the processing of subject–object ambiguities in both Dutch and German. This research has by-and-large confirmed Frazier's (1987) initial finding of a subject-before-object preference. Representative work that has contributed to establish the SO preference includes Hemforth(1993), Schriefers et al. (1995), Konieczny (1996), Bader and Meng (1999), Scheepers et al. (2000), Schlesewsky et al. (2000). A comprehensive review can be found in Bader and Bayer (2006).

The study of Lamers (2001, 2005) is one of the few studies that compares the use of case marking to other sources of disambiguating information. She investigated

[1] The original formulation of the AFH in Frazier (1987) was phrased in terms of Augmented Transition Networks (ATN), a grammar formalism popular in computational linguistics during the 1970s (cf. Allen 1987).

subject–object ambiguity resolution in Dutch main clauses like those in (9) in an experiment measuring event-related potentials.

(9) *De oude vrouw in de straat verzorgde* ***hem/hij*** *...*
the old woman in the street took-care-of him/he ...
'The old woman in the street took care of him' / 'He took care of the old woman in the street ...'

Due to the general SO preference, people start to interpret the initial NP as the subject of the sentence. Encountering the accusative case-marked pronoun *hem* fits this initial parse; if the nominative case-marked *hij* is encountered, it turns out that the initial NP is not the subject, but the object of the sentence, and a garden-path effect was accordingly observed. To investigate whether these processes were directly related to the use of case information, Lamers tested also sentences in which the animacy of the initial NP in combination with the selectional restriction of the verb provided the disambiguating information, as exemplified in (10). Strikingly, processing of sentences starting with an inanimate NP did not differ from processing of sentences with an initial animate NP.

(10) *De oude vrouw/ Het oude park in de straat* ***verzorgde*** *hij ...*
the old woman the old park in the street took-care-of he ...
'He took care of the old woman / the old park in the street ...'

More recently, Lamers and de Hoop (2005) analysed both studies in the newly developed framework of incremental optimal interpretation. This model applies principles of optimality theory in an incremental, word by word fashion (see also de Hoop and Lamers 2006, and de Hoop Chapter 6). The analysis clearly showed that the same constraint satisfaction pattern was observed, and that in both object-initial sentences the language user 'jumps' from the initially optimal subject-initial interpretation to the less preferred object-initial interpretation. Already at the verb the word-order constraint PRECEDENCE ('The subject precedes the object') is violated for sentences starting with an inanimate NP, whereas CASE ('The subject is assigned nominative and the object accusative') is satisfied. The same is true at the case-marked pronoun *hij* in (9). This is illustrated in the tableaux 26.1 and 26.2.

Tableau 26.1. Incremental optimization of the object initial sentence in (9)

De oude vrouw ... 'the old lady'	verzorgde ... 'took care of'	hij ... 'he'	case	precedence
☞ Subject Initial	☞ Subject Initial	Subject Initial	*	
Object Initial	Object Initial	☞ Object Initial		*

978–0–19–920647–6 27-Spencer-ch26 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 410 of 418 May 16, 2008 18:59

Tableau 26.2. Incremental optimization of the inanimate initial sentence in (10)

Het oude park… 'the old park'	verzorgde… 'took care of'	animacy	precedence
☞ Subject Initial	Subject Initial	*	
Object Initial	☞ Object Initial		*

Combining psycholinguistic data with a theoretical OT perspective, the incremental optimization of interpretation analysis shows how the use of case information relates to other constraints in language, and thereby provides one way to bridge the gap between psycholinguistics and theoretical linguistics.

Returning to the issue of the AFH, note that this principle only applies to active fillers, that is, fillers which can be unambiguously identified as such either by their position or by their lexical form. This raises the question of what happens with fillers that are not active in the given sense, as, for example, in so-called scrambling sentences like (11). In (11a) with SO word order, both subject and object of the embedded clause are in their canonical position. Under widely, although by no means universally shared assumptions, the OS word order in (11b) comes about by scrambling the object in front of the subject, leaving a trace behind.

(11) a. *Ich weiß nicht, ob* [*die Frau*] *einige Lehrer angerufen* ***hat.***
I know not if the woman some teachers called has
'I do not know if the woman called some teachers.'

b. *Ich weiß nicht, ob* [*die Frau*]$_i$ *einige Lehrer* t_i *angerufen* ***haben.***
I know not if the woman some teachers called have
'I do not know if the teachers called the woman.'

Given the lack of a filler–gap dependency in (11a), the phrase *die Frau* ('the woman') is not an active filler, and the AFH therefore cannot apply. However, another principle proposed for filler–gap processing could apply. Based on evidence from Italian, De Vincenzi (1991) generalized the AFH to the Minimal Chain Principle (MCP) shown in (12).

(12) Minimal Chain Principle (MCP)
Avoid postulating unnecessary chain members at S-structure, but do not delay required chain members.
(De Vincenzi 1991: 13)

The Minimal Chain Principle predicts a preference for SO in sentences like (11). This prediction has been borne out in many experiments (e.g. Bader and Meng 1999; Friederici and Mecklinger 1996). In fact, locally ambiguous sentences in which

978–0–19–920647–6 27-Spencer-ch26 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 411 of 418 May 16, 2008 18:59

an object has been scrambled in front of the subject are among the most difficult garden-path sentences found in German.

From the perspective of the HSPM, the major problem posed by subject–object ambiguities is the correct assignment of syntactic functions to each phrase of a sentence. Prima facie, this seems to be a problem genuinely involving the processing of case information. However, as shown by the discussion thus far, the problem of identifying syntactic functions in simple clauses has been rephrased as a special case of the more general problem of processing filler–gap dependencies. While filler–gap dependencies are often seen to arise from syntactic movement, this is not necessarily so. Different theories of syntax have offered various devices to capture the intuition behind filler–gap dependencies, namely that some word or phrase does not appear in its canonical position, including devices which do not involve syntactic movement or not even gaps in a literal sense (e.g. Bresnan 2001; Gazdar et al. 1985; Pollard and Sag 1994).

However, even if one subscribes to the more narrow view that chains are a device that can be used by the grammars of natural languages, and thus that filler–gap dependencies are actually dependencies between a trace and a moved phrase (= the trace's antecedent), the problem of assigning syntactic functions still does not reduce completely to the problem of correctly locating traces. Several instances of syntactic-function ambiguities have been discussed in the psycholinguistic literature for which the alternative syntactic structures do not differ in terms of filler–gap dependencies. A first example for such an ambiguity is given in (13).

(13) a. *(Alle) Menschen, die in Not sind, sollte man unterstützen.*
all people who in distress are should one support
'One should support (all) people who are in distress.'
b. *(Allen) Menschen, die in Not sind, sollte man helfen.*
all people who in distress are should one help
'One should help (all) people who are in distress.'

Sentence (13a) contains a verb taking an accusative object, whereas the verb in (13b) subcategorizes for a dative object. This difference is visible when the sentences in (13) contain the bracketed determiner: accusative *alle Menschen* versus dative *allen Menschen*. However, if the initial NP is used as a determinerless bare NP, a local ambiguity arises which is only resolved when the sentence-final verb makes it clear what case the initial NP bears.

Hopf et al. (1998; see also Hopf et al. 2003) measured event-related potentials in order to investigate how readers process the local ambiguity between accusative and dative case. Despite the fact that this ambiguity involves only a tiny single feature, namely the case feature of the initial NP, Hopf et al. found clear evidence for a garden-path effect when sentences were disambiguated by a verb requiring a dative object, as in (13b). The ambiguity in (13) can be considered a 'pure case ambiguity' which is not reflected in the phrase-structure representation. Thus, readers seem

to decide on the case of a case-ambiguous NP in the absence of sufficient evidence even if this decision is not accompanied by any phrase-structural consequences.

A similar point can be made about a certain subclass of sentences which are locally ambiguous between an SO- and an OS-structure without involving any filler–gap dependencies at all. A relevant sentence pair is provided in (14).

(14) a. *(Keiner wußte,) daß Maria ein Schnitzel spendiert hat.*
Nobody knew that Maria[NOM] a schnitzel[ACC] bought has
'Nobody knew that Maria bought a schnitzel.'

b. *(Keiner wußte,) daß Maria ein Schnitzel spendiert wurde.*
Nobody knew that Maria[DAT] a Schnitzel[NOM] bought was
'Nobody knew that a schnitzel was bought for Maria.'

At least under parsimonious syntactic assumptions, the two sentences in (14) are both base-generated. For (14a), this is so because this is a simple sentence exhibiting SO word order. (14b) is a representative of a subclass of OS sentences for which OS is the canonical, base-generated word order. This subclass contains sentences with passivized ditransitive verbs, as in (14b), as well as sentences with experiencer-object psych-verbs and different types of unaccusative verbs. Often, the object in these sentences is animate and the subject inanimate, and this might be a reason for such sentences having OS instead of SO as base-order (cf. Haider 1993).

Nevertheless, a robust generalization concerning such sentences is that they exhibit a preference for SO order, thus conforming to the general pattern described above (cf. Bader and Bayer 2006; Meng and Bader 2000). Given that no filler–gap dependencies are involved, neither the Active Filler Hypothesis nor the Minimal Chain Principle applies. In order to account for preferences in favour of certain cases – and thereby syntactic function assignments – even in the absence of associated phrase-structural differences, Hopf et al. (1998) postulated the Case Preference Principles shown in (15), which are stated under the assumption that nominative and accusative are structural cases in German whereas dative is a lexical (or oblique) case.

(15) Case Preference Principles
a. Prefer structural Case to lexical Case.
b. Prefer nominative Case to accusative Case.

The Case Preference Principles are rather surface oriented. A question which is outside the scope of the current chapter is whether the Case Preference Principles can be subsumed under more general principles (for further discussion, cf. Bader and Bayer, 2006).

To summarize, psycholinguistic studies on syntactic function ambiguities in Dutch and German have revealed important information about the role that case

information might play in sentence comprehension. Depending on the position of the sentence and its function in relation to the ambiguity of the information, different kinds of processes can be elicited varying from lexical processing to reanalysis. Stated in terms of case, the robust generalization has emerged that nominative case is preferred to accusative case (true for both Dutch and German), and accusative case in turn is preferred to dative (applies only to German).

The task of assigning the correct syntactic function to a phrase can be considered a prime example of a task involving the processing of case features, at least in languages in which syntactic functions are not rigidly associated with specific phrase-structural position. Nevertheless, as pointed out repeatedly in the preceding discussion, specifying exactly how the HSPM makes use of case information meets several obstacles. Thus, while the generalizations summarized above are robust in empirical terms, what they show with respect to the HSPM's processing of case information is less so. A major reason for this is the fact that there is no consensus with regard to the proper place of case within the mental grammar. In our opinion, progress in this area will only be possible by a closer collaboration between theoretical linguists and psycholinguists working on the phenomenon of case (see also Lamers and Ruigendijk, this volume, Chapter 27.)

Before we can close the topic of syntactic function ambiguities, a final point has to be made. So far, we have discussed the resolution of syntactic function ambiguities in purely linguistic terms. The parsing principles that we have cited as possible candidates for guiding the HSPM's decisions in the face of syntactic ambiguity all make reference to structural notions, including phrase-structural configurations and morphosyntactic features. An alternative basis for deciding between competing syntactic structures is provided by probabilistic information derived from frequency of usage. The question of whether the HSPM's actions are governed by structural decision principles or by decision principles based on frequency (or by a mixture of both) is at the centre of current psycholinguistic research. A general review of the role that is played by frequency information for human on-line language processing is provided by Jurafsky (2003).

With regard to the processing of syntactic function ambiguities, this question has not been decisively settled yet. Overall, sentences with SO word order are clearly the majority, as expected given that SO is the canonical word order in Dutch or German (for an early corpus-based study showing this for German, see Hoberg 1981). This however does not mean that the finding of an SO-preference for these languages is reducible to frequency information. As a matter of fact, it might well be that certain word orders are canonical because they have privileged properties with regard to information structure (cf. Vallduvi and Engdahl 1996), which in turn might explain simultaneously why they are preferred in situations of ambiguity and why they occur more frequently than other word orders.

Bornkessel et al. (2002) have presented evidence from an ERP study which they claim to show that at least certain aspects of computing syntactic functions is

independent from corpus frequencies, but the validity of their corpus analysis has been questioned by Kempen and Harbusch (2002).

26.4 Using Case Information to Identify Clauses

We now turn to the processing of case in strictly head-final languages like Japanese and Korean (for an overview of topics in Japanese language processing, cf. Mazuka and Nagai 1995). In these strictly head-final languages, the second ambiguity problem identified in section 26.2 – the problem of identifying clause boundaries – is much more pressing than in languages like English, Dutch, or German. The problem arises because the left-side of embedded clauses is neither signalled by complementizers nor by specifically marked elements like relative pronouns or wh-words. A sentence fragment like the one in (16) is therefore multiply ambiguous (this and the following examples are taken from Inoue and Fodor 1995).

(16) *Bob-ga* *Mary-ni* *ringo-o* . . .
Bob-NOM Mary-DAT apple-ACC

Two possible continuations of (16) are shown in (17).

(17) a. [$_S$ *Bob-ga* *Mary-ni* *ringo-o* ***ageta.***]
Bob-NOM Mary-DAT apple-ACC gave
'Bob gave Mary the apple.'
b. [$_S$ *Bob-ga* *Mary-ni* [$_S$ *ringo-o* **tabeta**] ***inu-o*** ***ageta***]
Bob-NOM Mary-DAT apple-ACC ate dog-ACC gave
'Bob gave Mary the dog that ate the apple.'

In (17a), (16) is continued with the three-place verb *ageta* ('gave'). This verb can take the three NPs of (16) as its arguments and can thus complete (16) as a main clause. In (17b), in contrast, the next word following the three NPs of (16) is the two-place verb *tabeta* ('ate'). Since *tabeta* does not take a dative argument, an embedded clause must be postulated on encountering this verb in (17b). Given the following accusative NP *inu-o* ('dog'), this embedded clause turns out to be a relative clause with the prior accusative NP *ringo-o* ('apple') as its object. According to Inoue and Fodor (1995), readers show a mild surprise effect when encountering *tabeta* in (17b), suggesting that the three NPs of (16) are initially integrated into a single clause which is only split up into two clauses when no other option remains.

While the small garden-path effect observed for (17b) is practically unavoidable given that a sequence of a nominative NP, a dative NP, and an accusative NP

978–0–19–920647–6 27-Spencer-ch26 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 415 of 418 May 16, 2008 18:59

perfectly matches, the requirements of a ditransitive verb, other sequences of case-marked NPs might cause the language comprehender to anticipate an upcoming embedded clause early on. An example illustrating this point is given in (18).

(18) [$_S$ *Bob-ga* *Mary-ni* [$_S$ *ringo-ga* *atatta*] *inu-o* *ageta*]
Bob-NOM Mary-DAT apple-NOM hit dog-ACC gave
'Bob gave Mary the dog that the apple hit.'

The first two NPs of (18) are identical to those of (16). The third NP, *ringo-ga* ('apple'), however, is now a nominative-marked NP. Since there is normally only one nominative NP per clause,[2] the nominative-marking on *ringo-ga* is a strong clue that a clause boundary has to be inserted in front of this NP. Taking case information into account is thus potentially of great help to the HSPM.

As an example experiment looking at the role of case information for identifying clause boundaries in Japanese and Korean, we will consider an experiment by Kim (1999). Sentence (19a) contains a locally ambiguous centre-embedded relative clause. Sentence (19b) is a control sentence with coordinated VPs.

(19) a. Relative clause Condition
Wunchensoo-(ka/nun) [[*chunksoboo-lul* *sealduk-han*] *gunchukka-lul*]
driver-(NOM/TOP) janitor-ACC persuade-REL architect-ACC
bibanhanda.
criticized
'The driver criticized the architect who persuaded the janitor.'
b. Coordination Condition
Wunchensoo-(ka/nun) [*chunksoboo-lul* *sealduk-hago* *gunchukka-lul*
driver-(NOM/TOP) janitor-ACC persuade-COOR architect-ACC
bibanhanda].
criticized
'The driver persuaded the janitor and criticized the architect.'

The initial NP was either marked with the nominative particle *–ka* or with the topic particle *-nun*. The question addressed by Kim was what happens when the HSPM encounters the disambiguating second accusative NP *gunchukka-lul*. When encountering the preceding verb *sealduk-han*, the HSPM will need to construct an embedded clause (either a relative clause or a complement clause). Is the initial NP put into this embedded clause, or does it stay in the main clause, and what effect does the particle have on this decision? Because topic-marked NPs are more likely part of a main clause than nominative-marked NPs, Kim's prediction was that processing the disambiguating NP in sentence 26.5a) should be easier with a topic-marked initial NP than with a nominative-marked initial NP. This is what

[2] There are some well-known exceptions to the one-to-one correspondence between nominative (*ga*-marked) NPs and clauses which we will not consider here.

Kim actually found in a self-paced reading experiment. For sentences with VP-coordination, in contrast, no reading time differences depending on the particle were found.

In summary, the HSPM's decisions concerning clause boundaries seem indeed to be sensitive to information associated with case particles in languages of the Japanese/Korean type. Both verb subcategorization information, as in (18), and information-structural properties, as in 26.5), are crucially involved in these decisions.

26.5 Case and Markedness

Above, we reported a preference for accusative case over dative case for locally ambiguous German sentences. Additional experimental results show that dative case plays a special role not only with regard to syntactic ambiguity resolution, but also for several other processing phenomena.

One set of findings concerns ungrammatical sentences as in (20). These have been derived from the locally ambiguous sentences in (14) by replacing the case-ambiguous proper name *Maria* by NPs marked for case.

(20) a. *(*Keiner wußte,*) *daß der Mutter ein Schnitzel*
Nobody knew that the[DAT] mother a schnitzel-[ACC]
spendiert hat.
bought has

b. *(*Keiner wußte,*) *daß die Mutter ein Schnitzel*
Nobody knew that the[NOM] mother a Schnitzel-[NOM]
spendiert wurde.
bought was

Sentence (20a) is ungrammatical because the initial NP *der Mutter* is a dative NP instead of a nominative NP. The reverse is true for sentence (20b). Here, the initial NP *die Mutter* is a nominative NP instead of a dative NP.[3] A robust finding concerning these ungrammatical sentences is that in experiments in which sentences have to be judged as grammatical or not under time pressure (speeded grammaticality judgments), the ungrammaticality inherent in (20a) is reliably detected by native speakers whereas the ungrammaticality inherent in (20b) is often overlooked (cf. Bader and Bayer 2006; Meng and Bader 2000). Several pieces of evidence show

[3] The two NPs *der Mutter* and *die Mutter* are in fact both case-ambiguous in ways which are immaterial to the issue at hand.

that the erroneous acceptance of ungrammatical sentences like (20b) is a performance effect. For example, errors only occur when the distance between the initial NP and the clause-final auxiliary is lengthened, for example by inserting adverbial material in between. Further findings include the phenomenon of case attraction (cf. Bader and Bayer 2006) and differential effects of garden-path strength (cf. Meng and Bader 2000).

Somewhat more abstractly, the ungrammaticality in (20a) can be described as one where the HSPM has to assign nominative case to a dative NP; in (20b), in contrast, dative case has to be assigned to a nominative NP. Under the assumption that dative case is a marked case, the experimental results for such sentences can be described by saying that under tight processing conditions the HSPM is unwilling to override a marked case by an unmarked case (cf. 20a) whereas an unmarked case is easily overwritten by a marked case (cf. 20b). The assumption that dative case is generally marked in German is a controversial assumption, however. Some authors have argued that the dative case assigned by ditransitive verbs is an unmarked, structural case (e.g. Wegener 1990; Wunderlich 1997). Since most of the experimental findings cited in this section have been obtained with sentences containing ditransitive verbs, finding a special role of dative case in these circumstances argues in favour of the assumption that dative case is a marked case throughout.

If so, the findings from German give rise to the hypothesis in (21).

(21) Markedness in language comprehension
The HSPM does not assign marked case unless forced by unambiguous input. Once assigned, the HSPM is unwilling to withdraw a marked case.

A more thorough discussion of this hypothesis, together with syntactic analysis of the markedness relations inherent in the German case system, can be in found in Bayer et al. (2001) To our knowledge, German is the only language for which markedness within the case system has been systematically explored with respect to its processing by the HSPM. Future research will show whether similar findings will be obtained for other languages, especially for languages with richer case systems and case systems defined in different ways.

26.6 Summary

This chapter has reviewed psycholinguistic work on the role of case for human syntactic parsing. One recurrent theme of this chapter has been that a substantial amount of work relevant in this context is only indirectly concerned with case. For example, a large body of research concerned with syntactic-function ambiguities

had as its prime focus the processing of filler–gap dependencies, with the assignment of case being a mere side effect of the position of the gap associated with the filler. However, as also discussed in this section, preferences for particular case assignments do not seem to be completely reducible to parsing decisions concerning phrase-structure parsing. Together with the findings on processing effects related to markedness that were reviewed in section 26.5, this suggests that at least certain aspects of case play an independent role within the HSPM. Despite all progress, it seems fair to say that research on the processing of case is still in its infancy. In particular, the set of languages that have been investigated is still severely limited. Going beyond this narrow language base will be of prime importance for making progress in our understanding of how case is processed by the HSPM.

Unfortunately, reasons of space prevented us from discussing some recent developments related to the processing of case. For example, important evidence on the question of the relationship between thematic information and case has been adduced by Bornkessel and colleagues (Bornkessel et al. 2003, 2004; Bornkessel and Schlesewsky 2006). In addition, the topics of contextual influences and auditory sentence comprehension have been addressed in work by Knoeferle and Crocker (2006).

CHAPTER 27

CASE IN APHASIA

MONIQUE LAMERS
ESTHER RUIGENDIJK

27.1 Introduction

As is clear from previous chapters, case provides the language user with a rich source of information for conveying meaningful messages in natural language use. Psycholinguistic research has shown that in the *intact* human language system, morphologically realized case might help to distinguish the arguments, identify the syntactic function, and trigger certain parsing mechanisms, such as attachment decisions, establishing phrase boundaries, and, thus, providing essential information for comprehension (see also Bader and Lamers, this volume, Chapter 26). The question arises if and how a disturbance of the language system as in aphasia affects this ability to use case information. Aphasic patients suffer from neurological damage to the language system resulting in deviant patterns of language use. Whereas some patients seem to exhibit profound expressive deficits often involving syntactic impairments, others seem to suffer from severe language comprehension problems. In the next section it will be discussed that according to the classical approach (see Geschwind 1965) in aphasiology these two roughly sketched language profiles characterize two main aphasic syndromes each of which is associated with a specific neuroanatomical basis: Broca's and Wernicke's aphasia.

We start with a brief introduction to aphasiology sketching the language problems known to characterize the different aphasic syndromes. We will discuss some

important accounts for the aphasic impairment in language comprehension either directly addressing morphological case in the impaired language system, or highly related issues, such as the comprehension of canonical and non-canonical sentences. For production, the emphasis will be on cross-linguistic findings on the realization of case-marked entities such as DPs and pronouns. Differences between structural and lexical case will briefly be addressed as well. Additionally, the use of case in comprehension and production will be related to the most important explanations that have been suggested for the language impairments in aphasic syndromes. We conclude with a summary and some directions for further research.

27.2 On Aphasia

Aphasia is a language disorder that is caused by brain damage, such as a cerebrovascular accident (CVA) or a traumatic brain injury. It can affect all language modalities, speech production, auditory comprehension, reading and writing, at all linguistic levels: phonology, morphology, semantics, and/or syntax. The type and severity of the language impairment vary from person to person, but nevertheless some syndromes that share several symptoms can be distinguished. Because of their contrastive language profiles, the focus in this chapter will be on the two major syndromes Broca's and Wernicke's aphasia.

Broca's aphasia is characterized by non-fluent telegraphic output, which is also known as agrammatic speech, consisting mainly of content words and a lack of function words, with relatively intact speech comprehension. It is normally caused by a lesion in the third frontal convolution of the left hemisphere, the subcortical white matter, extending posteriorly to the inferior portion of the precentral gyrus (the motor strip) (Goodglass and Kaplan 1983). These areas are also known as Broca's area (Figure 27.1), named after Paul Broca, one of the first researchers who associated language deficits in aphasia to damage of specific brain areas.

Mainly based on studies in English and Italian, Caramazza and Berndt (1985: 33) defined agrammatic speech or agrammatism as a morphosyntactic impairment, with the following main characteristics: 'the omission of grammatical morphemes, reduced phrase length, the omission or nominalization of verbs, and difficulties with word order'. More recently, Menn et al. (1995) give a detailed description of agrammatism incorporating language-specific differences. They state that agrammatic patients rely to a great extent on the simplest possible structures of their language. Especially in languages like English, Dutch, Italian, or German this results

978–0–19–920647–6 28-Spencer-ch27 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 421 of 436 May 16, 2008 18:59



Figure 27.1. Left Hemisphere with Broca's area in the premotor cortex and Wernicke's area in the superior temporal cortex (adapted from Greenfield 2001: 162)

in the omission of function words, such as auxiliary verbs, prepositions, pronouns, and articles. Bound grammatical morphemes are less likely to be omitted, but may be substituted, and verbs appear with the most common inflection or in a form that does not require agreement (the infinitive). A speech sample of a Broca aphasic is given in Box 27.1.

Box 27.1 Example of agrammatic speech production

I had a stroke. Blood . . . low. low blood pressure. Period. Ah . . . pass out. Uh . . . Rosa and I, and . . . friends of mine . . . uh . . . uh shore uh drink, talk, pass out *and when you woke up, where were you*? Hahnemann Hospital . . . uh uh I wife, Rosa . . . uh take . . . uh love. ladies . . . uh ocean uh hospital an'transfer Hahnemann hospital ambulance uh half 'n hour . . . uh uh it's . . . uh motion, motion . . . uh bad patient . . . I uf flat on the black. Um It's uh . . . shaved, shaved. Nurse shaved me, uh shaved me, nurse, uh wheelchair . . .

From Menn and Obler (1990)

Whereas comprehension of *words* is relatively spared in Broca's aphasia, comprehension of *sentences* is generally assumed to cause specific problems, especially with sentences that deviate from the canonical word order. They show more problems with the interpretation of passives, object clefts, and object relatives, than with

actives, subject clefts, and subject relatives.[1] Apart from these problems, the interpretation of pronouns is impaired as well (see Grodzinsky et al. 1993; Ruigendijk et al. 2006; Vasić et al. 2006 for details).

Wernicke's aphasia is usually caused by a lesion in the posterior portion of the first temporal gyrus of the left hemisphere. In the late 1870, it was Carl Wernicke who linked this brain area, also known as Wernicke's area, with a rather fluent language profile with problems in both language production and comprehension at the word and the sentence level. The fluently articulated speech is characterized by many paraphasias (i.e. unintended production of phonemes, words, or phrases, see Goodglass and Kaplan 1983) and may be paragrammatic. Paragrammatism is characterized by long and complex sentences, doubling sentence parts or blending two sentences, and substitution errors with grammatical morphemes rather than omissions (Huber et al. 1983). In Box 27.2, the English-translated speech sample illustrates the paragrammatic output of a Wernicke's aphasic.

Box 27.2 Example of speech production from a Wernicke's aphasic speaker

Here in Munich there said he has he me has he me, he has the eye, made the eye made. There said always: my name is so and so XXX yes, is even so nice when you are named not? Yes my name is how he said my name so and so, and then said I Mr Professor, how. For me, really but for me not, what should I with him?

From Ruigendijk (2002, translated from German)

(XXX indicates unintelligible speech)

Historically, these two aphasic syndromes are classified as being fully distinctive from each other, forming a dichotomy on the neuroanatomical level with anterior lesions for Broca's and posterior lesions for Wernicke's aphasia, and on the functional level with disabilities that relate mostly to syntactic processes for Broca's, and for Wernicke's aphasia with lexical/semantic disabilities. Lesion studies including patients of both populations are of particular interest for the investigation of case in natural language, since case encompasses syntactic as well as semantic information. Over the years, it has become clear that the distinction is not as strict as was assumed in the classical approach. On the one hand, several studies have shown a selective comprehension deficit in Broca's aphasics with problems in comprehension of non-canonical reversible sentences (cf. Grodzinsky 2000), while, on the other hand, the language profile of some Broca's aphasics seems to lack the agrammatic characteristics (e.g. Saffran et al. 1989). Moreover, if the speech of Broca's aphasics

[1] For a discussion on the exact nature of the comprehension problems, see Grodzinsky (2000), and replies to this paper, as well as several articles in issues of *Brain and Language* (1999–2006).

is characterized as being agrammatic, this does not mean that these patients have no syntactic awareness at all (cf. Kolk and van Grunsven 1985; Kolk and Friederici 1985; Hartsuiker et al. 1999a). Furthermore, Wernicke's aphasics may, apart from lexical-semantic deficits, also show grammatical difficulties (cf. Heeschen 1980). Nevertheless, in clinical linguistics the two syndromes are still taken as diagnostically distinguishable[2] on the functional level and on the neuroanatomical level.[3] In the next sections, relevant sentence comprehension and production studies will be discussed.

27.3 Case morphology in sentence comprehension

In this section, we will review studies that addressed the question if and how aphasic speakers use morphological case marking for the interpretation of non-canonical sentences (e.g. passive and object-relative sentences). In agrammatism, the interpretation of these sentences has often been shown to be problematic compared to canonical sentences.[4] Moreover, research, initiated by a study by Caramazza and Zurif (1976) showed that agrammatic aphasics made more errors in the interpretation of semantically *reversible* object relatives like (1a) than in semantically *irreversible* object relatives as in (1b):

(1) a. The cat_i that the dog is chasing t_i is black.
b. The $house_i$ that the man is painting t_i is blue.

[2] For an extensive discussion on the variability within and between Broca's and Wernicke's aphasia see Drai and Grodzinsky (2006).

[3] Nowadays, modern techniques such as positron emission tomography (PET), and functional magnetic resonance imaging data (fMRI), play a major role in the precise assignment of language functions to brain areas (Indefrey and Levelt 2004). There are, however, no neuroimaging studies that specifically addressed the use of case information in the impaired language system. Therefore, the information that will be provided on brain areas in relation to different language functions and disabilities in this chapter will be mainly based on clinical impressions and assumptions.

[4] Not all lesion studies report impaired comprehension of Broca's aphasics on these types of sentences. For example, in the study of Blumstein et al. (1998) Broca's aphasics did not differ from normal participants on several filler–gap constructions (wh-questions, subject and object relative clauses, embedded wh-questions). Moreover, Berndt et al. (1996), who performed a meta-analysis on published studies concerned with the comprehension of reversible sentences, report mixed results in Broca's aphasics. However, in a recent study using event- related brain potentials in Dutch clear effects are reported, also providing evidence for the use of semantic information in irreversible sentences (Wassenaar 2005)

According to several linguistic theories, in both sentences the first DP has been moved out of its base-generated position behind the verb and leaving a trace t_i (or copy, depending on the theory). Psycholinguistic studies have shown that during on-line processing the moved element, i.e. *the cat* in (1a), is reactivated at the trace position (e.g. Garnsey et al. 1989; Kaan et al. 2004). One of the structural theories providing an account for the deficit in agrammatism with filler–gap dependencies is the trace deletion hypothesis (Grodzinsky 1990, 2000; Grodzinsky and Finkel 1998). According to this account, traces are deleted from the syntactic representation in agrammatism, therefore theta-roles cannot be transferred, which results in comprehension problems of sentences in which the proper assignment of theta-roles depends on the presence of traces of moved elements, such as non-canonical sentences (passives, object-relatives etc.). In irreversible sentences, semantic cues can be used to derive a meaningful interpretation with correctly assigned theta-roles, whereas in irreversible sentences such cues are unavailable (see Wassenaar 2005).

In languages in which arguments are morphologically case-marked, case marking might be used to assign thematic roles correctly. German, for example, is a language in which DPs are overtly marked for case (see (2) for an example). This case-marking is mainly realized on the determiner and/or adjective. It provides information on the syntactic function, which, according to Primus (Chapter 17, this volume) in combination with other structural information sources, might lead to the correct thematic role assignment and thus correct interpretation (see also Bader and Lamers, this volume, Chapter 26). Unfortunately, the number of lesion studies on languages with morphological case marking is relatively small, and, as we will see, the results are not straightforward. We first present some studies on German, followed by some data from other languages.

Von Stockert and Bader (1976) found an effect of case morphology in German Wernicke's aphasics, in that their performance was better when case morphology could be used for sentence interpretation. This effect was not found for Broca's aphasics. The authors argued that the agrammatic patients relied more on lexical and logical information than on morphological cues. However, this finding could not be replicated in a later study by De Bleser et al. (1988), who showed that sensitivity to case morphology in agrammatic speakers was at least preserved for simple SVO structures. Heeschen (1980), who used a sentence–picture matching task with semantically irreversible and reversible sentences, all unambiguously marked by morphological means, reported similar results. He found that Broca's aphasics were more sensitive to case morphology than Wernicke's aphasics, although they also used a semantic strategy taking plausibility into account to comprehend the sentences. This study was followed up in German by Burchert et al. (2001, 2003) and de Bleser et al. (2005), who analysed the comprehension of various types of canonical and non-canonical sentences with and without disambiguating case morphology by Broca's aphasics more systematically. Some examples are given in (2).

(2) a. *der Junge sucht den Vater* (active, SVO)
the.NOM boy is seeking the.ACC father
b. *den Vater sucht der Junge* (active, OVS)
the.ACC father is seeking the.NOM boy
c. *die Frau sucht das Kind*
the.NOM woman is seeking the.ACC child
the.ACC woman is seeking the.NOM child
(active, ambiguous between SVO and OVS)
d. *der Vater wird von dem Jungen gesucht* (passive)
the.NOM father is being by the.DAT boy.DAT sought

The agrammatic subjects in general could discriminate different case markings (free and bound morphemes) on the word level (Burchert et al. 2003, de Bleser et al. 2005). This did not mean, however, that they could use case morphology to improve sentence comprehension of non-canonical sentences. The interpretation of ambiguous case-marked sentences (such as in 2c) was compared to that of unambiguously case-marked sentences (e.g. 2a, 2b). The agrammatic speakers did not show a difference between the ambiguous and unambiguous sentences, and thus as a group[5] were not helped by case morphology for their sentence interpretation.

There are some studies on Serbo-Croatian, a language in which case is expressed as a suffix on the noun. Smith and Mimica (1984) showed that Serbo-Croatian-speaking Broca's aphasics were severely impaired in the use of case morphology for sentence comprehension, but not in the use of semantic information. Replicating this study, Smith and Bates (1987) compared the use of case and gender morphology and found similar results for the use of case morphology; however, patients were more sensitive to gender morphology, although this was impaired too. These authors found that when three cues (word order, gender, and case) were presented together, the performance on non-canonical sentences by Broca's aphasics was close to normal.

Lukatela and colleagues (1988) concluded, based on results with a grammaticality judgment task in which the correctness of oblique case (accusative, dative) had to be judged, that their Serbo-Croatian agrammatic patients were still sensitive to inflectional morphology, since they rejected 88 per cent of the ungrammatical sentences.

Hagiwara and Caplan (1990) showed a canonicity effect for Japanese, in which case-marking particles accompany the arguments. They examined semantically reversible sentences in canonical and non-canonical word order and found that case marking is not enough to interpret the sentences correctly. Unfortunately, the participants in this study were not classified into aphasia types, which makes it difficult to interpret the results.

[5] Note however, that there was quite some individual variation.

Finally, Friedmann and Shapiro (2003) tested Hebrew agrammatic aphasic speakers on the interpretation of active sentences with a canonical word order (SVO) and a non-canonical word order (OVS and OSV). In Hebrew, the object is marked unambiguously for case. The results showed that the aphasic speakers were sensitive to the existence of the case marker, but they were unable to use the case morphology to compensate the deficit in the interpretation of non-canonical sentences.[6]

27.4 Theoretical implications of the use of case in aphasic comprehension

The studies discussed so far show that the performance of Broca's aphasics in relation to the use of case information is very diverse. Besides the reported differences between different groups of aphasics in different studies, type of task seems to be of influence on the aphasics' performance. Nevertheless, it can be concluded that agrammatic aphasics recognize case-marking but have difficulties using it efficiently in the comprehension process. Cross-linguistically, in most studies agrammatic patients do not seem to benefit from the presence of case information alone (see Friedmann and Shapiro 2003; Burchert and others 2003; De Bleser et al. 2005).

Whether the diverse patterns of the use of case-marking in the comprehension of reversible sentences can be looked upon as support for the trace deletion hypothesis as a feasible account for agrammatism remains unclear. Initially one might think that the *inability* of Broca's aphasics to use case information for comprehension supports this hypothesis. After all, according to the trace deletion hypothesis Broca's aphasics have a reduced syntactic representation of the structure. Although case marking can be recognized, for which a trace is not necessary, it is impossible for the aphasics to use this information as a cue to understand the sentence and derive the base-generated position.

It can, however, not be ruled out that the case effect results from processes that are fully independent of structure derivational processes such as trace detection. The difference in morphological form might simply contribute to the distinguishability of the two arguments even without being able to link the arguments to the base-generated position, but still facilitating the comprehension process (de Hoop and Lamers 2006). Unfortunately, the studies under

[6] The results showed a chance performance on OVS and OSV. If they had completely ignored the case marker they would have performed below chance and treated the sentences as SVO.

978–0–19–920647–6 28-Spencer-ch27 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 427 of 436 May 16, 2008 18:59

discussion do not provide any information on the exact nature of underlying processes.

Thus, even though the case effect can be explained as being independent of the trace deletion hypothesis, there seem to be no clear arguments against it directly stemming from the agrammatic speakers' use of case information. There are several (other) reasons why the trace deletion hypothesis has been criticized. Among others the trace deletion hypothesis seems to fail to provide an explanation for the discrepancy between on-line comprehension and high performance on grammaticality judgment tasks, as found by Linebarger et al. (1983), but more recently also by Wassenaar (2005). Some authors argue that the trace deletion hypothesis cannot deal with individual differences in severity of the impairment (e.g. Caramazza et al. 2000, Berndt et al. 1996; but see replies by Grodzinsky in 2000 or Drai and Grodzinsky 2006), whereas experimental results indicate such a difference (Wassenaar 2005).

Besides the trace deletion hypothesis, there is another structural hypothesis that may account for the agrammatic comprehension problems, namely the mapping hypothesis (Linebarger et al. 1983), following which the source of the agrammatic deficit is not the extraction of the syntactic information, but rather the mapping of the syntactic information onto the semantic representation, the syntax–semantic interface. As a consequence, the comprehension of non-reversible sentences is disturbed because syntactic information, including morphological case marking, derived from the input cannot be mapped onto the semantic representation. If sufficient semantic information is available, such as animacy information in reversible sentences, the mapping can come about solely on the basis of semantic information.

The mapping hypothesis has also been challenged. For example, Swinney and Zurif (1995) found an early syntactic breakdown in the processing of relative clause sentences for Broca's aphasics, whereas the data of Vigliocco and Zilli (1999) show the ability of Broca's aphasics to map conceptual as well as syntactic gender information onto the representation. Moreover, the mapping hypothesis is not very specific in what exactly causes the breakdown in the mapping and whether this is an early, or later state in sentence comprehension (e.g. the ability to check the subcategorization frame of the whole sentence for gaps and fillers or the actually link between a possible filler to a gap).

Another explanation for agrammatism, as well as other aphasic language deficits, is of a more general cognitive nature, possibly involving disruptions or limitations in normal use of functions that are not per se specific for language comprehension, such as working memory. According to this theory, the problems that are observed in agrammatic Broca's aphasia result from a limitation in processing capacity, which affects computational efficiency (see e.g. Kolk 1995; Piñango and Burkhardt 2005; Avrutin 2006). This limited capacity theory is one of several theories that associate different types of errors in the different aphasic syndromes with the same

underlying deficit (see Butterworth and Howard 1987, Bates et al. 1991). This deficit is either directly reflected in the language production (which is argued to be the case in Wernicke's aphasia with its erroneous and paragrammatic output) or is only indirectly shown, because of an adaptation to the deficit (i.e. in agrammatic output; see Kolk 1995 for a detailed account along these lines). Recently, the limited processing capacity idea has been elaborated. Piñango and Burkhardt (2005) and Avrutin (2006) among others, propose that the limitation in processing capacity affects syntactic processing specifically, which is slowed down, or weakened in agrammatic Broca's aphasia (Slow-Syntax resp. Weak Syntax Hypothesis). Avrutin suggests that since syntax is not the most automatic linguistic module for the interpretation and production of sentences anymore, other linguistic routes, such as discourse operations, may 'take over'. This results in aberrant interpretation of non-canonical sentences and pronouns, but also in the omission of determiners, pronouns, and verb inflection from speech production (Piñango 2002; Avrutin 2006). To our knowledge there are no specific suggestions in this approach with respect to the ability, or the lack of ability of agrammatic speakers to use (case) morphology.

27.5 Case in production

Although the production of case morphology in aphasia (or the lack thereof) is mentioned in several studies, only a few address this issue specifically. Nevertheless, we will try to give as exhaustive an overview as possible of what is known about case in aphasic production in this section. As for comprehension, most research concentrates on agrammatism.

Already in 1905, Kleist observed the '… almost exclusive use of the nominative and accusative case and the infinitive and participle forms with deletion of function words' in a German aphasic speaker (1905: 504), which indicates that at least some German aphasics may have problems with the realization of at least dative and genitive case. Bayer and colleagues (1987) followed up on this early observation and examined the production of case morphology with a structured task, in which German patients with Broca's (n = 3) and Wernicke's aphasia's (n = 2) had to insert a determiner and adjectival case marking in different types of sentences (such as sentences with canonical (subject–verb–object) and non-canonical order (object–verb–subject), double object sentences, accusative and dative case in PPs and genitive case in possessive NPs). This study was partly replicated by De Bleser

and others (2005) with seven more agrammatic speakers. The agrammatic speakers were able to produce nominative and accusative in simple SVO sentences, but they performed poorly on dative objects. Note however that some of the control subjects made many errors on these items also (the authors report scores of 55 to 67 per cent correct for dative sentences). Some of the agrammatic speakers appeared to use some kind of strategy, either using nominative most of the time, or accusative. Other patients seemed to apply a word order strategy: nominative comes first and accusative second. The Wernicke's patients made many incongruency errors between determiner and adjective. The agrammatic patients that Bayer and others (1987) reported were unable to use correct case morphology in double object sentences, but case in PPs and in possessive constructions was relatively unimpaired (around 75–80 per cent correct). The Wernicke's patients showed the same pattern, but with more problems in possessive constructions. They not only made errors with case, but also with gender. The authors concluded that agrammatic aphasics specifically have problems with the correct use of case morphology in syntactic relationships with a wider scope (such as double object sentences).

Ruigendijk (2002) suggested that case assignment as such is not impaired in agrammatic aphasic production, but that categories that depend on case may be problematic as a result of the agrammatic problems with case-assigning elements. One of the key characteristics of agrammatic aphasia is the problems these patients have with verbs and verb finiteness (see e.g. Menn and Obler 1990; Saffran et al. 1989). Exactly these elements are crucial for case assignment. The hypothesis is that the realization of complete and correctly case-marked DPs in agrammatic speech is related to the realization of case assigning categories. When case cannot be assigned, since the relevant category is missing, a DP may be realized as a bare NP or with default (usually nominative) case. This hypothesis was tested using spontaneous speech analyses and several elicitation tasks with Dutch, German, and Russian aphasic speakers (see Ruigendijk et al. 1999 and Ruigendijk and Bastiaanse 2002 for details on experiments and results; see also Ruigendijk and Friedmann 2002 and Ruigendijk 2002 for Hebrew and Hungarian). Generally, these studies confirmed the hypothesis that more complete and correct DPs (nouns with a determiner or pronouns) were used in agrammatism when a case assigner was present and that very often the default case was used (nominative in these languages) when no case assigner was present. This meant that for subject DPs a finite verb was present, and for object DPs a verb. Prepositions also counted as case assigners. Only those instances where it could be clearly established whether or not a case assigner was present were analysed. Noun phrases for which no case assigner was present were either subjects in combination with a non-finite verb or noun phrases in utterances in which no verb was realized at all.

We will provide more detailed information on each language below.

27.5.1 Dutch

Dutch has no overt case marking on the noun phrase anymore, apart from some fixed archaic expressions. Dutch pronouns still have different subject and object forms. Ruigendijk and others (1999) tested the hypothesis that the presence of a complete DP in agrammatic production depends on the realization of a case assigning category. This study involved a spontaneous speech analysis with Dutch and German data in which for every noun and pronoun it was established whether a case assigner was present or not.[7] The data confirmed the hypothesis: when the case assigner was absent, virtually no accusative and dative case-marked object pronouns were produced, but only nominative pronouns, as well as noun phrases without a determiner.[8] When a case assigner was present, the noun phrase was realized more often with than without a determiner. These findings were confirmed in another study by Ruigendijk (2002) examining sentence production data from Dutch agrammatic speakers.

27.5.2 German

In the same line of experiments, Ruigendijk (2002) described results from ten German agrammatic speakers (see also Ruigendijk and Bastiaanse 2002). They were examined with sentence completion and production tasks. Their spontaneous speech production was also analysed. The results confirmed the hypothesis: agrammatic speakers realized significantly more complete noun phrases than incomplete noun phrases (i.e. the determiner was omitted) in the presence of a case-assigning category, both in spontaneous speech and in the experimental tasks. When no case assigner was realized, either the determiner was omitted, or nominative (which is default case in German) was realized. The results on the experimental tasks demonstrated that the patients performed relatively well on the production of nominative subject DPs. The performance on accusative and dative objects assigned case by a verb or by a preposition was significantly impaired. This also held for genitive case assigned by a preposition or as an adnominal genitive. Most errors were 'case substitution errors', and most of these reflected a tendency to overgeneralize one case. Seven patients overused accusative case for objects and two overused dative for objects. Crucially, the use of case morphology improved when the case-assigning category was realized.

The same tasks and the spontaneous speech analysis were also done with one German Wernicke's aphasic patient. Her spontaneous speech data demonstrated

[7] Since this is sometimes difficult to examine in agrammatic production, not all DPs were included in the analysis.

[8] Note that for Dutch noun phrases only the completeness could be established, that is, whether or not a determiner was realized. Dutch pronouns were also analysed with respect to their overt case marking.

no problems with case marking; she hardly ever omitted or substituted determiners. Interestingly, her performance on the tasks did not differ from that of the agrammatic speakers. She had hardly any problems with nominative and accusative case assignment, whereas dative elicited more errors.

27.5.3 Russian

Tsvetkova and Glozman (1975) analysed the spontaneous speech of aphasics and reported the frequent omission of verbs and pronouns as well as problems with word order. Case substitution errors were found also, in that nominative case (default in Russian) was used instead of other cases. Luria (1976) also described the speech production of agrammatic-speaking patients. He reported a low number of verbs, a decrease in the number of non-nominative cases and an increase in the number of nouns with nominative case (see also Tsvetkova 1969 and Ryabova 1970). One patient frequently omitted the verb of the sentence and realized dependent nouns with nominative instead of accusative or dative case on a sentence repetition task. Akhutina (1991) analysed the speech production of agrammatic speakers in successive stages of recovery and found that patients started out with pure nominalizations (nominative), which developed in structures with nouns and verbs. Initially, patients were only able to distinguish between nominative and accusative case, whereas in later stages genitive and prepositional case could be used as well. Dative and instrumental case appeared even later.

Ruigendijk (2002) reported data from seven Russian agrammatic speakers, who were examined with the same setup as described for her studies in Dutch and German. The results of the two tasks and the speech analysis showed that the Russian agrammatic speakers produced more correctly case-marked DPs than incorrectly case-marked DPs when a case-assigning category (i.e. depending on the structural position for instance a finite verb, a verb, or a preposition) was present. As in German and Dutch, when no case assigner was present, default nominative case was preferably used in spontaneous speech production. For the sentence completion task, the participants were shown a picture, which depicted an action and an incomplete sentence. The sentence had to be completed with the missing DP, normally being marked with either nominative, accusative, dative, or instrumental case. The agrammatic patients performed worse on the realization of all four cases than the non-brain-damaged controls. Most errors involved case substitution. Dative case was the most difficult for these patients, followed by instrumental case. Accusative and nominative case caused fewer problems. In the sentence production task, in which the participants had to describe a picture in one sentence, the agrammatic speakers seemed to prefer the use of default nominative case instead of the other cases.

27.5.4 Hungarian

McaWhinney and Osmán-Sági (1991) tested nine Hungarian speakers with Broca's aphasia and five with Wernicke's aphasia with a sentence production task, and found that both Broca's and Wernicke's aphasics omitted nominative DPs frequently, which reflects a common process in Hungarian: subject ellipsis. Broca's aphasics omitted the indirect object quite often as well, whereas the Wernicke's aphasics omitted the direct object. Kiss (2000) found for two agrammatic-speaking patients on a sentence production task that their responses often consisted of isolated arguments for which no (case-assigning) verb was realized. Most of these isolated arguments were realized without case marking, which is nominative case in Hungarian. This is in line with the results described above for German and Russian. According to Kiss (2000), the tendency to omit case marking indicated that the agrammatic patients were impaired in syntactic structure-building operations and could not create the minimally required domain needed for the case assignment procedure.

Ruigendijk (2002) presents results from a study with one Hungarian agrammatic aphasic patient. A spontaneous speech analysis showed that this patient did not make any case-marking errors on DPs when a case-assigning verb was present. Isolated DPs for which no case-assigning verb was realized and that could not be analysed as an elliptical utterance were produced most often as nominatives. The patient made relatively few errors on elicitation tasks. He performed equally well on all case types, with one exception: he found the production of sublative DPs (as in (3)) that were assigned lexical case by a verb almost impossible.

(3) *A nőnek pénz hiányzik a pénztárcájából. A gyermek új labdát vett.*
A nő gyanakszik . . . *completion:* a gyermekre
The woman lacks money from her wallet. The child has bought a new ball.
The woman suspects . . . *completion:* the child.SUBL

27.5.5 Hebrew

Ruigendijk and Friedmann (2002; see also Ruigendijk 2002) analysed the speech production of seven Hebrew-speaking agrammatic aphasic speakers with regard to their ability to realize the accusative case marker *et*, which must be realized with definite objects and which cannot be used with indefinite objects. These patients produced case-marked objects only if a case-assigning verb was realized as well. Interestingly, they were sensitive to the relationship between definiteness and the presence of the case marker. Only three times, they produced a definite object

without the case marker (out of 122 DPs in total) and only three times an indefinite object with an illicit accusative marker.

27.6 Theoretical implications of the use of case in aphasic production

Ruigendijk (2002) and Ruigendijk and Friedmann (2002) relate the case-marking production abilities of the agrammatic aphasic speakers to their problems with syntactic structure building as proposed in the Tree Pruning Hypothesis (TPH, e.g. Friedmann 2001, 2006), as well as approaches in which the specific syntactic impairment results from a limited processing capacity (as in Piñango and Burkhardt 2005; Avrutin 2006). The TPH states that agrammatic speakers cannot project the syntactic tree to the highest nodes. The higher the node, the more problematic it is. Moreover, it suggests that if an agrammatic speaker cannot access or project a certain node (e.g. TP), s/he will also be impaired in structures that are related to higher syntactic nodes (e.g. relative clauses that depend on CP). Lower nodes are relatively spared. If syntactic structure building fails, and as a result the aphasic speaker cannot realize the proper case-assigning element, then case cannot be assigned, resulting in a DP without overt case marking or realized in the default nominative form. This implicates that the production of case morphology as such is not impaired in agrammatic aphasia.[9]

Although this seems to be true for spontaneous speech production, a dichotomy was found between the realization of structural case and lexical/inherent case in the production experiments in German, Russian, and Hungarian. Recall that the realization of structural case (i.e. nominative and accusative case) was relatively unproblematic, but lexical/inherent case assignment was not. Although not many case errors are found in spontaneous speech, when elicited experimentally, overall lexical cases seem to be more problematic than structural cases, even when the case-assigning element is realized. The difference may be explained by the fact that once the syntactic structure is built, structural case can be assigned automatically; it comes for free once the proper configuration is realized. This does not hold for lexical case. Even if a case assigner (say the verb) is realized, the lexical information that specifies lexical case assignment must be retrieved and this information must be used on time during syntactic encoding. Note that the limited capacity theory,

[9] For a more elaborate discussion on different accounts for how morphological case is realized in production the reader is referred to Melinger, Pechmann, and Pappert, Chapter 25, this volume.

978–0–19–920647–6 28-Spencer-ch27 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 434 of 436 May 16, 2008 18:59

discussed in section 27.4, also refers to possible timing problems in the syntactic encoding. Thus, problems retrieving or using information for lexical case assignment *on time* also explains some of the substitution errors from the German speakers who tend to overuse accusative case for objects, when lexical dative case should have been used.

27.7 Summary and Conclusion

To gain more insight in the role of case in the impaired language system an overview of deletion studies that more or less explicitly address this issue was presented. Unfortunately, the number of studies including Wernicke's aphasics is rather limited, making it difficult to compare the use of case information between Broca's and Wernicke's aphasia. Moreover, from the studies described here, it becomes clear that the investigation of the brain-damaged system is not a trivial matter. The findings are very diverse, especially for comprehension, leading to inconclusive results, making it hard to come to any robust generalizations on the use of case in the impaired language system. Nevertheless, in this paragraph, we will draw some conclusions, and point out some directions for further research.

From the studies addressing the use of case in the comprehension of reversible sentences, it can be stated that generally agrammatic aphasics are able to recognize morphological case marking in these sentences, but they are unable to use this case information for comprehension. We discussed several possibilities why agrammatic aphasics are not able to use information that in principle helps to distinguish the two arguments and can be accessed, namely a reduced syntactic representation, impaired mapping between the syntactic and semantic representations, and a limitation in processing capacity. Whereas the first account seems to be mainly syntactic in nature, the latter two accounts also incorporate more semantically related mechanisms. Unfortunately, the studies discussed in this chapter do not provide enough information to disentangle these two aspects, nor are we aware of any studies that do so.

For production, cross-linguistic studies have shown impairment in morphological case realization in agrammatic aphasics. Spontaneous speech production is mainly characterized by frequent omission of case morphology or resorting to default (nominative) forms. The finding that different patterns are observed depending on the presence of the case-assigning element is taken as evidence for impairment in structure-building processes, rather than in morphological case realization itself. This also explains the observed strategies, assigning nominative or accusative case as a default if dative case was required (see also section 26.5 in

978–0–19–920647–6 28-Spencer-ch27 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 435 of 436 May 16, 2008 18:59

this book). The sparse data on Wernicke's aphasia, the aphasic syndrome which is primarily characterized by a lexical-semantic deficit, suggest that there is at least some overlap in the performance of Wernicke's and agrammatic aphasics, but of course no strong conclusions can be drawn on this relatively limited amount of data.

Whereas the use of case information in language production seems to be mainly syntactic in nature, in comprehension the involvement of impaired semantic mechanisms could not be ruled out. To challenge this, studies are necessary that specifically and systematically address semantic and syntactic aspects of the use of case information. Moreover, several lesion studies have shown that by making use of modern research technologies (such as ERPs and magnetoencephalogram [MEG], used for neuroimaging), the exact time course can be tracked down and diverse mechanisms can be distinguished yielding different effects for diverse processes during language comprehension (e.g. Hagoort et al. 1996, Kolk et al. 2003, Kotz et al. 2003; Wassenaar 2005). However, none of these studies specifically address the use of case information, nor are there any specific studies on case using techniques that are especially suitable to track down the involved neuroanatomical areas. Especially, behavioural studies in combination with studies using these modern techniques will help us to gain more insight into the impaired language system, as well as in the function of case as an important linguistic phenomenon for both comprehension and production.

PART V

AREAL AND DIACHRONIC ISSUES

CHAPTER 28

EVOLUTION OF CASE SYSTEMS

LEONID KULIKOV

This chapter (largely based on Kulikov 2006) offers a cross-linguistic survey of the main types of possible developments in case systems. Section 28.1 focuses on the main mechanisms of the rise of new cases and expansion of case systems (case-increasing). New cases may arise (i) by adding adverbs, postpositions, and (rarely) prepositions (see section 28.1.1); (ii) by adding existing case-markers to other case forms, which results in 'multilayer' case marking (see 28.1.2); (iii) from demonstrative pronouns or articles (see 28.1.4). New case forms may also go back to (iv) denominal adjectives and adverbials incorporated into the case paradigm (see 28.1.3). An important mechanism of the rise of new case(s) is (v) splitting of one case into two by borrowing of a new case marker from a different declension type (see 28.1.5).

Section 28.2 discusses the main processes within case systems that do not lead to quantitative changes but help to resist phonetic erosion (stable case systems). The mechanisms used to avoid merger of cases include the borrowing of new inflections from other cases and adding free morphemes to old case forms.

On the basis of this diachronic typological overview, Section 28.3 offers a tentative classification of the evolutionary types of languages and briefly discusses the main factors determining the evolutionary type of a language.

28.1 Rise of cases and expansion of case systems: Sources and mechanisms

28.1.1 Case morphemes from adpositions

Most often, new case-markers (and, accordingly, new cases) are recruited from adpositions. Particularly common are case suffixes originating from postpositions or other semi-auxiliary adverbial words with similar semantics.[1] Thus, in Harris and Campbell's (1995: 89) formulation, 'Cases develop from postpositions when the postposition is felt to be so closely connected to its attribute noun that together they are reinterpreted as one word; semantic and morphophonemic changes (e.g. vowel harmony) often take place which conceal the word boundary and change the status of elements, resulting in new case suffixes.' The origin of case prefixes from prepositions, albeit theoretically possible, seems to be extremely rare in the languages of the world (see some examples in section 28.1.1.3).[2] For the adpositional origin of case markers see also Kahr 1976; Lehmann 1995: 80ff.; Blake 2001: 161ff. This, the commonest scenario, can be illustrated by examples from the history of the Indo-Aryan,[3] Tocharian, and Lithuanian case systems.

28.1.1.1 *New cases from postpositions and nominal compounds in Indo-Aryan and Tocharian*

By the end of the Middle Indo-Aryan (MIA) period, that is, at the turn of the second millennium AD, the Indo-Aryan languages have lost most of the cases of the original Sanskrit, or Old Indo-Aryan (OIA), system of eight cases[4] (which, except for minor details, is nearly identical to the case system reconstructed for Proto-Indo-European). Generally, only two cases survive, Direct (resulting from the merger of nominative and accusative) and Oblique (mostly going back to the Old Indo-Aryan genitive), although in some languages isolated traces of some other oblique

[1] The adpositions, in turn, often go back to verbal forms or (especially often for expressing locative case relations) to nominal forms; see, for instance, Blake 2001: 161ff.

[2] This disproportion may be due to some reasons of general nature. Thus, as Reh (1986) concludes on the basis of the analysis of evidence from some African languages, prepositions may tend to become suffixes on the preceding verb rather than prefixes on the following noun. This, in turn, according to Stampe, may result from the fact that '[u]naccented elements tend to attach rather permanently to preceding accented elements, and not to following ones' (Donegan and Stampe 1983: 344; Stampe 1994).

[3] Case morphemes going back to postpositions can also be found in many languages of another branch of the Indo-European language family, closely related to Indo-Aryan, in Iranian; cf. a few examples from Ossetic in section 28.1.3, fn. 9.

[4] For a general survey and discussion of the evolution of the Indo-Aryan case system, see, in particular, Bloch 1934; Zograf 1976; Bubeník 1998: 99–101; Bubeník and Hewson 2006: 102ff.; Masica 1991: 230ff.

978–0–19–920647–6 29-Spencer-ch28 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 441 of 457 May 16, 2008 18:59

cases, such as instrumental, locative, or ablative, can still be found, sometimes even within the declension paradigm; cf. the Sinhala instrumental case suffix *-en/-in* and Assamese ergative *-e*, both reflecting the OIA instrumental singular ending of *a*-stems *-ena*. The functions of the lost cases are largely taken over by morphemes (bound or free, i.e. postfixes or postpositions) of different origin.

These include:

(i) Primary, or 'old', postpositions, going back to Proto-Indo-European (PIE) morphemes that were used in the adpositional function already in the proto-language. An example of an old OIA postposition reflected as a case suffix in a daughter language, in Middle Indo-Aryan, is the Māhārāùñrī ablative suffix *-āhi* < Skt. postposition *adhi* (constructed with ablative in OIA); see Insler 1991–92; Bubeník 1998: 68f.

Next to old postpositions, there are several markers which result from grammaticalization of some verbal and nominal forms:

(ii) Postpositions descendent from
  (ii.a) non-finite verbal forms, in particular, converbs (traditionally called 'absolutives', or 'gerunds', cf. Skt. *ādāya* 'with', lit. 'having taken'), gerundives (participia necessitatis) and verbal adjectives (cf. (3));
  (ii.b) case forms of some nouns (cf. (1–2));

(iii) final members of compounds, which, again, may represent
  (iii.a) non-finite verbal forms (cf. nominal compounds in *-sthita-* mentioned in (1a)) or
  (iii.b) nominal case forms (cf. nominal compounds in *-artham* 'goal, purpose' in (2)).

The markers of the first three types (i, ii.a and ii.b), representing free morphemes (words), were originally attached to (non-nominative) case forms of the noun, thus forming, after having become bound morphemes, the second layer of case forms (see below). In type (iii), the source of the new case morpheme was attached to the nominal stem, thus creating a new case within the first layer. In fact, due to the erosion of the nominal inflection by the end of the MIA period, some (oblique) case forms may eventually become indistinguishable from bare stems, and thus the border between types (ii) and (iii) cannot always be drawn with accuracy.

A number of examples of the grammaticalization of new postpositions and case suffixes can be found already in the MIA period, in particular, in Apabhraüśa Prakrits (for details, see Bubeník 1998: 67, 80), cf. the ablative postfix *-ṭṭhiu* < OIA *sthita-* 'standing' (passive perfect participle of the verb *sthā* 'stand') and the locative postposition *majjhe* < Skt. *madhye* (loc.sg. of *madhya* 'middle') in (1):

(1) Apabhraṃśa Prakrit (Bubeník 1998: 67, 80; Bubeník and Hewson 2006: 113)
  a. *hiaya-ṭṭhiu*
    heart-LOC
    'out of [my] heart'

978–0–19–920647–6 29-Spencer-ch28 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 442 of 457 May 16, 2008 18:59

b. *gharaho majjhe*
house.GEN in
'in the house'

In New Indo-Aryan languages we observe a rapid increase in use of such new postpositions, which are normally added to the Oblique case form. This grammaticalization may result in the amalgamation of a postposition with the nominal stem or Oblique case and, hence, in the rise of a new case. Such is, for instance, the origin of some new case endings listed under (2):

(2) a. Sinhala dat. *-ṭa*, Khowar dat. *-te* < Skt. *-artham* 'goal, purpose';
b. Sinhala gen. *-ge* < Skt. *gr̥he* 'in the house' (Loc.sg. of *gr̥ha-* 'house')

The bulk of case markers containing *k-* and/or *r-*, which go back to nominal derivatives of the Old Indo-Aryan verbal root *kr̥-* (*kar-*) 'make, do', can be found in several New Indo-Aryan languages. These include, in particular, genitive morphemes in several New Indo-Aryan languages (see, in particular, Bubeník and Hewson 2006: 122f.):

(3) New Indo-Aryan genitive
a. Hindi *-kā*, *-ke* < Apabhr. *-kera* < Skt. gerundive *kārya-* 'to be done';
b. Awadhi, Maithili *-ker* < Skt. part.pf.pass. *kr̥ta-* 'done, made';
c. Bhojpuri *-kæ* < Skt. adj. *kr̥tya-* 'to be done'.

Likewise, some dative *k*-morphemes, such as Hindi *-ko*, Oriya *-ku*, Marathi *-kē*, Romani *-ke*/*-ge* reveal a vestige of the same Sanskrit root *kr̥-* (*kar-*).

The initial stages of the corresponding grammaticalization processes can be dated as early as Old Indo-Aryan. Thus, the starting point of the grammaticalization path of Skt. *-artham* 'goal, purpose' towards the Sinhala dative case suffix *-ṭa* (cf. (2)) is the adverbial usage of the accusative of the Sanskrit bahuvrīhi compounds in *-artha-* (*X-artham*), meaning 'having X as a goal, purpose' → 'for (the sake of) X': *udakārtham* 'having water as a goal' → 'for water'; *sukhārtham* 'having happiness, pleasure as a goal' → 'for happiness, for pleasure'; *tadartham* 'having that as a goal' → 'for that, therefore'.

In New Indo-Aryan languages, the morphological status of the resulting markers may vary from bound morphemes (case suffixes), tightly connected with the nominal stem (as in Sinhala, cf. (2)), to free morphemes (postpositions). The latter type can be illustrated by the Hindi dative–accusative morpheme *-ko* which can be shared in some constructions by several nouns (as in *rām aur mohan ko* 'to Ram and Mohan'), exemplifying a 'Gruppenflexion', which pleads for a postposition rather than for a suffix analysis.

The difference between these groups of case morphemes is often described in terms of the distinction between cases of the first, second, and third layers (Zograf 1976; Masica 1991: 230ff.; Matras 1997). The first layer corresponds to the case in

the strict sense of the term and, in Hindi, is limited to the opposition between the direct and oblique cases. The third layer corresponds to clear instances of postpositional phrases, while the second one takes an intermediary position between cases proper and postpositional phrases. It is important to note that only the first layer case can trigger agreement on adjectives. Although both 'Gruppenflexion' and the lack of agreement with second layer cases appear to distinguish these morphemes from cases proper, the high degree of grammaticalization makes it appropriate to associate them with the category of case in general.

The genesis of the Tocharian case system (for a detailed survey, see, in particular, van Windekens 1979; Pinault 1989: 71ff.; Bubeník & Hewson 2006: 317ff.) resembles in several respects the origins of the Indo-Aryan cases. Proto-Tocharian (the proto-language of Tocharian A and B) has lost about a half of the original Proto-Indo-European system. Alongside with the nominative and oblique (which continues the Proto-Indo-European accusative), we only find continuations of the genitive and vocative (only in Tocharian B) and, probably, some traces of the ablative. Nevertheless, the paradigm has even more cases than in Proto-Indo-European. Like in New Indo-Aryan languages, next to the inherited, or 'primary', cases, nominative, oblique, genitive, and vocative, there are a number of 'secondary' cases built on the oblique form. Historically, most of them go back to combinations with postpositions, although the exact sources of some endings are unclear. Thus, the locative morpheme Toch. B *-ne*, Toch. A *-aṃ* reflects PToch. **-næ* < PIE **no*, thus being probably related to the Old Prussian locative preposition and prefix *na* (< **no*) 'on'; cf. also Slavic *na* (< **no-H*) 'on'. The perlative[5] morpheme (Toch. B *-ā*) must continue PToch. **-a(C)* and thus may be related to Latin *ad*. Still less clear is the origin of the allative morpheme, Toch. A *-ac*, Toch. B *-ś(c)* < PToch. **-cä*, which presumably originates from **-Te* or **-Ti* (where *T* stands for any dental), thus reflecting **-de* (cf. Greek δε), **-dhi* (cf. Greek θι), or **-te* (cf. Greek -τε).[6]

28.1.1.2 *New locative cases from postpositions in Old Lithuanian*

Similar mechanisms for creating new cases have been used in Old Lithuanian, resulting in a subsystem of locative cases, which is partly preserved in the modern literary language[7] and can still be found in some archaic dialects, spoken, in particular, in Belarus (for details, see e.g. Ambrazas et al. 1985: 90 [= 1997: 106]; Mathiassen 1996: 38; Zinkevičius 1996: 112f.; Seržants 2004; Kortlandt 2005; Bubeník and Hewson 2006: 206ff.). Most of the Lithuanian cases can be traced back to the Proto-Indo-European case system. In addition, we find three new locatives: illative,

[5] This case takes over some functions of the instrumental.

[6] Alternatively, Toch. B *-ś(c)* can be explained as reflecting the PIE root *st(h)eH*$_2$ (*-sti-*) 'stand'; see Bubeník and Hewson 2006: 320.

[7] Some isolated forms still existing in modern Lithuanian do not belong to the paradigm but function as adverbials.

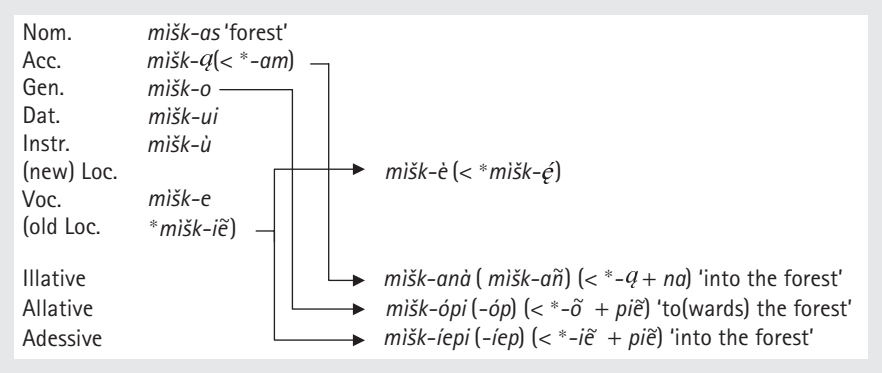


Figure 28.1. Old Lithuanian case system

adessive, and allative, made by attaching the postposition **nẵ* to the accusative form in *-n* (with the subsequent degemination of *-n-n-*) and the postposition **piẽ* (< Proto-Baltic **prei*) to the locative and genitive, respectively. The 'locative' of the modern literary Lithuanian case system (inessive) is a new case as well (made by adding the new postposition **en* to the old locative),[8] replacing the old locative inherited from Proto-Indo-European (which only survives in some adverbials such as *nam-iẽ* 'at home'). The resulting case system (as attested in Old Lithuanian and in some Southern and Eastern archaic dialects) is summarized in Figure 28.1.

28.1.1.3 *Case prefix from a preposition: Iranian and Nuristani languages*

Rare examples of the origin of a case prefix from a preposition are found, for instance, in a few Indo-Iranian (in particular, Pamir) languages, where the accusative case prefix goes back to a preposition. Thus, in Yazgulyam (Iranian, Pamir), the prefix *š(ə)-/ž(ə)-* (going back to an ablative preposition) is attached to the nouns preceded by the ablative preposition *na* 'from' (Comrie 1981b: 169; Payne 1980: 174). In Prasuni (Prasun), a (Nuristani) language spoken in Afghanistan, the form with the locative prefix *tu-* (*ti-*, *t-*) co-exists within the case paradigm with a number of suffixal forms, cf. (4):

(4) Prasuni (Morgenstierne 1949: 220)
  a. *ɛsl'ɛk tu-g'ul*
     that in-country
     'in that country'
  b. *t-arek*
     in-house
     'in(to) the house'

[8] For the history of this case form, see Kortlandt 2005: 68.

c. *tī-zī*
in-door
'at the door'

This case prefix has developed from the Common Indo-Iranian preposition *antár* 'within, inside, between' (Morgenstierne 1949: 220).

28.1.2 Multilayer case marking

New cases (usually, new locatives) can also be created by adding existing case-markers to some case forms or to adverbials with case-like semantics, which results in multilayer case marking.

Such mechanism has given rise to the rich system of locative cases in Finnish, as well as in some other Finno-Ugric (FU) languages (see e.g. Hakulinen 1961: 67ff.). Although the documented history of the FU languages goes back only about five hundred years and we cannot directly observe the rise of new cases (as in Indo-Aryan), most of the Finno-Ugric case-markers are morphologically transparent and their sources can readily be reconstructed. Thus, the three internal locative cases, inessive, elative, and illative, have been created by adding the case markers of essive and partitive (historically going back to locative–ablative) to the forms in *-s* (which can be identified as an adverbial suffix). Likewise, the external locatives (adessive, ablative, and allative) have been made from the adverbial forms ending in *-la*/*-lä*, which is probably identical with the final element in a few nouns with locative semantics such as *ete-lä* 'south', *pohjo-la* 'north', *appe-la* 'home of the father-in-law' (← *appi-* 'father-in-law'). The history of the Finnish declension is summarized on the basis of Hakulinen 1961 in Table 28.1:

Of course, it is sometimes nearly impossible to draw with accuracy the distinction between the two types discussed in Sections 28.1.1.1–2 and 28.1.2, i.e. between 'adpositional' cases and multilayer case marking. Thus, Proto-Lithuanian **nă* used in the formation of the illative (cf. *mišk-añ* < **-am* + *na* 'into the forest') might represent an old postposition or an extinct case-marker.

28.1.3 Cases from adjectives and adverbials

Adjectives and adverbials that may become incorporated into the substantive paradigm, represent another important source of new cases. This grammaticalization path has thus far received much less attention in the literature on the grammaticalization of cases than postpositions.

An instructive example is provided by Ossetic, an Indo-European (Iranian) language which has lost most of the Proto-Indo-European cases but, eventually, has developed an even larger case paradigm (for details, see Cheung forthcoming). Next

978–0–19–920647–6 29-Spencer-ch28 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 446 of 457 May 16, 2008 18:59

Table 28.1. The reconstructed history of the Finnish case morphemes

	Finnish	Proto-Finnic	Proto-Volga-Finnic	Proto-FU
Nom.	*-ø*	<		**-ø*
Acc.	*-n*	<		**-m*
Gen.	*-n*	<		**-ń*
Essive	*-na/-nä*	<		**-na*
Partitive	*-ta/-tä, -a/-ä*	<		**-δa* (*loc.-abl.)
Translative	*-ksi*	<	**-k-s-e*	
Internal locative cases				
Inessive	*-ssa/-ssä*	<	**-s-na/-s-nä*	
Elative	*-sta/-stä*	<	**-s-ta/-s-tä*	
Illative	*-hVn, -ñ, -seen*	< **-seń, -zeń*		
External locative cases				
Adessive	*-lla/-llä*	< **-l-na/-l-nä*		
Ablative	*-lta/-ltä*	< **-l-ta/-l-tä*		
Allative	*-lle(′)*	< **-leń*		
Abessive	*-tta/-ttä*	<	**-s-ta-k / -s-tä-k*	
Comitative	*-ine-*	<	**-i-n* (?)	

to two 'old' cases, nominative and genitive, directly continuing the corresponding Proto-Indo-European cases, as well as two new cases based on combinations with postpositions,[9] a few members of the case paradigm are probably of adjectival and adverbial origin. The comitative morpheme *-imæ* (only in the Iron dialect) must reflect an adverbial morpheme, cf. Avestan *maṯ* 'together, jointly'. Two other case forms are likely to go back to denominal adjectives incorporated into the substantive paradigm. The inessive ending may reflect the adjectival suffix **-īi̯a-* (cf. Vedic *párvata-* 'mountain' – *parvatī́ya-* 'growing in the mountains'). The equative morpheme probably originates in the adjectival suffix *-vant-* (as in Vedic *tva-* 'you' – *tvā́vant-* 'like you').

For the sake of convenience, the history of the Ossetic case paradigm is summarized in Table 28.2 (adopted from Cheung forthcoming), which represents the declension of the word *sær* 'head' in the singular; in the cases where the forms attested in two main dialects, Iron and Digoron, are different, the Iron form is given first.

Another instance of a new case marker of adjectival origin is found in Armenian. The gen./dat./abl.pl. ending *-c'* can be traced back to PIE **-sko-m*, where **-sko-* represents the derivational adjectival suffix (cf. Goth. *manna* 'man' – *mannisks* 'human', Old Church Slavonic *člověkъ* 'man' – *člověč-ьskъ* 'human; of the men').

[9] The dative ending *-æn* may go back to **ana* (cf. Avestan *ana* 'upon, over, across') or **anu* (cf. Old Persian *anuv*, Avestan *anu* 'along, after, according to'). The adessive marker *-bæl* (Digoron) undoubtedly originates in **upari* 'above, upon, on' (the lack of the labial stop in the adessive ending *-yl* attested in another dialect, Iron, may be due to the adaptation to the inessive morpheme *-y*).

Table 28.2. Declension of Ossetic *sær* 'head' in the singular

Nominative	*sær*	< nom. **-ah* and/or acc. **-am*
Genitive	*sær-y/sær-i*	< gen. **-ahi̯a*
Ablative	*sær-æj*	< abl. **-āt*/ins. **-ā* + encl. particle **i*
Dative	*sær-æn*	< postpos. **ana* or **anu*
Adessive	*sær-yl/sær-bæl*	< postpos. **upari*
Comitative	*sær-imæ*	< adverb *mat*
Inessive	*sær-y/sær-i*	< adj. suff. **-ïi̯a-*
Equative	*sær-aw*	< adj. suff. **-vant-*
Allative	*sær-mæ*	< dem.pron. (loc.) **ahmi* + dir.particle **ā*

Apparently, Armenian has introduced the possessive adjective into the substantive paradigm, replacing the old genitive plural; later this form was expanded to the dative and ablative (see Godel 1975: 106; Kortlandt 1984: 100 [= 2003: 47]).

28.1.4 Cases from pronouns and articles (indexicals)

New case-markers can also be recruited from the set of demonstrative pronouns or articles (which in turn typically go back to pronouns).

Such was probably the origin of some case endings in Kartvelian languages.[10] According to A. Šanidze, the marker of the Georgian nominative case *-i* (the 'subject–object' form, approximately corresponding to the absolutive of ergative languages) may go back to the postposed demonstrative pronoun *-igi* (through the stage of the definite article), which was probably reanalysed as *ig-i* 'that-NOM' (Boeder 1979: 471, note 20; 474, note 35). Similar may be the origin of the Laz and Megrelian (Mingrelian) 'narrative' case marker *-k* (for discussion, see Klimov 1962: 17f., 36f.). The Georgian ergative suffix *-ma/-m*, the marker of another case encoding the two main syntactic arguments, undoubtedly goes back to the demonstrative pronoun *man* 'that; he'; Old Georgian still preserves the more archaic form of this ending, *-man*, which is formally identical to the source morpheme (Klimov 1962: 55).

Likewise, some scholars suggest a similar origin of the ergative case marker in two North-West Caucasian (Abkhaz-Adyghe) languages, Kabardian (*-mə*) and Ubykh (*-n*). According to Kumaxov (1971: 43, 158; 1989: 31f.), both case morphemes may go back to the definite article/demonstrative pronoun 'this' (Kabardian *mə*, Ubykh *jəna*).

Finally, rich evidence for the development of ergative case markers from demonstrative and pronominal forms is provided by a number of Australian languages;

[10] I would like to thank V. Chirikba and Y. Testelets for having discussed with me the Caucasian data.



Figure 28.2. Development of demonstrative pronouns

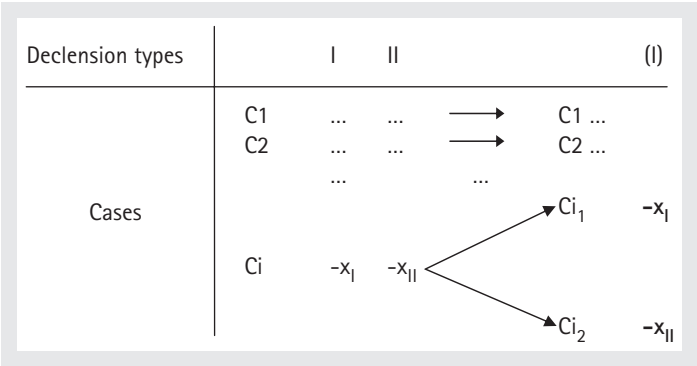


Figure 28.3. Case-splitting

see now the most comprehensive treatment of the issue in McGregor (to appear), which offers both a detailed survey of data available from several grammars and a theoretical explanation of this grammaticalization scenario. Thus, the Nyulnyulan ergative suffix *-ni/-in/-nim* etc. may derive from Proto-Nyulnyulan **-nimV*, which 'is a plausible cognate of the . . . third person pronoun *ni* ∼ *nu*'. The Jingulu ergative markers *-(r)ni* and *-nga* probably go back to either pronouns or demonstratives **-nu* and **-ngaya* (Ibid.).

In general, the possible developments of demonstrative pronouns can be represented according to the scheme in Figure 28.2. (Harris and Campbell 1995: 341f.)

28.1.5 Split of one case in two

New cases can also be created by splitting one case into two. The most common scenario is the borrowing of a new case marker from a different declension type, as shown in Figure 28.3. The case Ci that originally has distinct endings in declensions I and II ($-x_I$ and $-x_{II}$, respectively) splits, yielding two new cases, Ci_1 with the ending $-x_I$ and Ci_2 with the ending $-x_{II}$. This process is usually accompanied by the loss of the declension type (II) that was the source of the new case-marker.

An illustration of such development is provided by the history of the modern Russian 'second' locative and 'second' genitive (also called partitive).[11]

Locative-2 of the second declension type (historically going back to the Proto-Indo-European declension of the *-*o*-stems) is marked with the accented ending *-ú* (thus being opposed to *-e* of the original locative, or Locative-1[12]) and can only be employed in constructions with the locative prepositions *v* 'in' and *na* 'on'. It is formed almost exclusively from monosyllabic nouns such as *sneg* 'snow', *les* 'forest', *sad* 'garden': *na sneg-u*/*v sneg-u*, *v les-u*, *v sad-u*. The ending *-ú* has been borrowed from the old declension of the stems in *-*u*- (type *synъ* < **sūn-u-s* 'son', *dom-ъ* < **dom-u-s* 'house'), where it was regular: Nom.sg. *dom-ъ* (< **dom-u-s*) – Loc.sg. *dom-u* (< **dom-óu*). Under the influence of the locatives of some nouns of this type, such as *(v) med-u* '(in) honey', *-u*-forms have penetrated into the paradigm of the old *-*o*-type nouns, foremost of those which denote location and thus are particularly common in the locative usage. The earliest attestations of this new case appear at the turn of the thirteenth century. Subsequently, the *-*u*-declension has disappeared (approximately after the fourteenth century), being ousted by the productive second (*-*o*-) declension.

The genitive-2 (partitive) is distinct from the standard genitive (genitive-1) only for some uncountable nouns of the second declension, such as *mëd* 'honey', *saxar* 'sugar', *čaj* 'tea', cf. gen.-2 *mëd-u*, *saxar-u*, *čaj-u* ~ gen.-1 *mëd-a*, *saxar-a*, *čaj-a*. This form is employed in constructions such as 'a pound (of) ___', 'he drunk/ate some ___'. Like in the case of Locative-2, the ending *-u* has been taken over from the old *-*u*- declension, where it was regular: Nom.sg. *med-ъ* 'honey' – Gen.sg. *med-u* (< **medh-ou-s*); Nom.sg. *syn-ъ* 'son' – Gen.sg. *syn-u* (< **sūn-ou-s*, cf. Lith. Gen.sg. *sūnaũs*). Genitive-2 had been established approximately by the sixteenth century (for details, see e.g. Kiparsky 1967: 26ff.).

These developments are summarized in Figure 28.4, which represents the relevant fragment of the modern Russian second declension (in the singular) as compared to its Old Russian predecessor (attested since the eleventh century).

28.1.6 Other possible sources of new cases

Sections 28.1.1–5 do not exhaust the list of possible sources of new cases. Thus, a very interesting mechanism for the rise of a new case is instantiated by the emergence of

[11] For the synchronic status of these two cases, see, in particular, Zaliznjak 1967; Plungian 2002. For their history see e.g. Kiparsky 1967: 35ff.; Ivanov 1983: 266ff.; Hentschel 1991.

[12] Called in the Russian grammatical tradition *predložnyj*, i.e. 'Prepositional' case, since it can only be employed with prepositions.

Table 28.3. Old Russian nominative and vocative

	*-*ā*-declension	*-*o*-declension	*-*u*-declension
NOM.SG.	*žen-a* 'wife'	*vъlk-ъ* 'wolf'	*syn-ъ* 'son'
VOC.SG.	*žen-o* (< *g^wen-H_2-e)	*vъlč-e* (< *ulk^w-e)	*syn-u* (< *sūn-e-u)

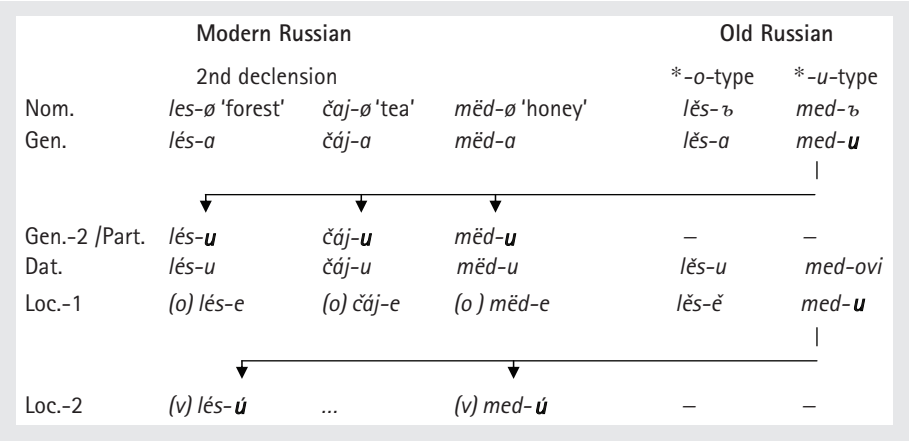


Figure 28.4. Development of the second declension in Russian

the new vocative in the modern (colloquial) Russian. The history of this form can be briefly summarized as follows.

Common (or Proto-) Slavic has inherited from Proto-Indo-European the old vocative that was distinct from the nominative. Originally, its form was equal to the bare stem or truncated stem + *-*e* (vocative particle?), but already by the Proto-Slavic period the morphemic structure has been blurred in all declension types, as shown in Table 28.3. Unlike all other Slavic languages, such as Polish, Bulgarian, and Ukrainian, modern Russian has not preserved this case form, which only survives in a few idiomatic expressions, such as *(o) bože!* 'Oh God!'. Instead, it has developed a new vocative form. The new vocative case is distinct from the nominative only for certain nouns of the first declension (historical *-*ā*-declension), specifically for those which are particularly common in the vocative usage. These include a few kinship nouns (*mama* 'mum(my)', *papa* 'dad', *tëtja* 'aunt') and short forms of proper names: *Maša* (short form for *Marija*), *Tanja* (∼ *Tat'jana*), *Saša* (∼ *Aleksandr* (m.)/*Aleksandra* (f.)), etc. The new vocative is marked with the zero ending, thus being identical with the bare stem (and with the genitive plural of this declension type): *mama – mam-ø*, *papa – pap-ø*, *tëtja – tët'-ø*, *Maša – Maš-ø*,

Tanja – Tan'-ø, *Saša – Saš-ø*, etc.[13] Most interestingly, the mechanism of such a resurrection of the vocative case, according to which the vocative form tends to be as close as possible to the bare stem (being in that respect similar to the imperative form in the verbal system), is likely to reflect a very basic (albeit of course not absolute and exceptionless) universal of the human speech.

28.2 Resisting the erosion of case morphology: Stable case systems

Section 28.1 illustrates the increase of the total number of cases, i.e. various quantitative changes in case systems; for examples of case syncretism and decrease of the total number of cases, see Chapters 14 and 30. Next to the case-increasing and case-decreasing types, we find more complex scenarios, where some developments compensate the loss of case oppositions, so that the number of case distinctions remains unchanged, but the system of forms may undergo crucial restructuring. The following section discusses some language mechanisms used for resisting potential case mergers.

28.2.1 Reinforcement of case forms

The most straightforward way to resist the erosion of the nominal inflection and, in particular, the loss of case distinctions, consists in the morphological reinforcement of case forms. The case morpheme can be reinforced by the repetition (redoubling) of the case marker (of the same or another declension type) and/or by adding an auxiliary word (for instance, a particle).

The former mechanism (repetition of the case morpheme) can be illustrated by the reconstructed history of the genitive form in Tocharian. According to van Windekens (1979: 181) and Pinault (1989: 88f.), the genitive ending *-ntse* in Tocharian B (< PToch. **-nsæ*) must represent PIE **-n-s-os*, i.e. the genitive of *n*-stems (with the ending *-s*), which was reinforced by the repetition of the same morpheme, when the form in **-ns* had become obsolete. Subsequently, this ending has been generalized for other stem types. The Tocharian genitive may thus instantiate an interesting example of the double (hypercharacterized) case marking.

[13] For an alternative analysis of this form as nominative (rather than bare stem), which underwent deletion of a prosodic unit ('deprosodization') with subsequent resyllabification, see Yadroff (1996). For the paradigmatic and morphological status of the vocative, see Chapter 43, this volume, and especially 43.2 on the status of the new Russian vocative.

A telling example of the latter type (adding an auxiliary word) is provided by Armenian,[14] which has preserved the original Proto-Indo-European system of case oppositions intact, in spite of the heavy phonetic erosion in the word-final position (auslaut), thus being even more 'case-stable' (see type III in section 28.3) than the phonologically more conservative Slavic and Baltic languages. Thus, the ablative singular forms (ending *-*os* in most declensions of the proto-language) have fallen together with the locative singular (*-*i*) after the apocope, i.e. after the loss of vowels in final syllables. For instance, both abl. and loc.sg. of the Proto-Indo-European word meaning 'heart' (*ḱr̥d-/*ḱērd(i)-; Proto-Armenian has generalized the stem *ḱērdi-) would yield Arm. *srti*. The removal of this syncretism has become possible by adding an enclitic particle (probably going back to PIE **eti*; reflected, for instance, in Skt. *áti* 'over, beyond'), which yielded *-ē*, cf. *sirt* 'heart' – loc.sg. *i srt-i* 'in the heart' ∼ abl.sg. *i srt-ē* 'from the heart' (Pedersen 1905: 221ff. [= 1982: 83ff.]; Kortlandt 1984: 103 [= 2003: 49f.]).

Two examples of case morphemes reinforced by particles are provided by the history of Ossetic declension (see Table 28.2). The allative ending *-mæ* is plausibly explained by Cheung (forthcoming) as based on the locative form of the demonstrative pronoun, **ahmi*, followed by the directional particle **ā*, first established in the pronominal paradigm and then expanded to the substantive declension. In the ablative, the ending *-æj* is likely to reflect both the old ablative (Proto-Indo-Iranian *-*āt*) and instrumental (Proto-Indo-Iranian *-*ā*), which have merged in the Proto-Ossetic *-*æ*, subsequently enlarged by the enclitic particle **i* (which is likely to have an ablative value).

28.2.2 Borrowing inflection from other cases: Genitive–accusative in Slavic

A more complex mechanism used to avoid merger of cases resulting from phonetic processes and erosion of inflection and thus to resolve possible case syncretism(s) and prevent the loss of cases[15] is the borrowing of new inflection from other cases.

One of the most well-known examples is the rise of the category of animacy in the history of Slavic languages (for details, see e.g. Klenin 1980; Iordanidi and Krys'ko 2000: 198ff.).

Already by the time of Common Slavic, the old Indo-European nominative and accusative had merged in most declension types due to the phonetic erosion of the end of the word (in auslaut): the final consonant was dropped, and the originally

[14] For a detailed discussion of all phonological changes and morphological developments which, altogether, result in a perfect preservation of the Proto-Indo-European system of case contrasts, see Meillet 1936: 64ff.; Godel 1975: 99ff.; Džaukjan 1982: 85ff.; Bubeník and Hewson 2006: 160ff.; and, especially, a short but very rich and insightful paper Kortlandt 1984.

[15] Cf. Harris and Campbell's (1995: 89) 'preservative' (or 'structure-preservative') reanalysis.

Table 28.4. Nominative, accusative, and genitive singular in Old Church Slavonic

	**-ā-*type *žena* 'wife'	**-o-*type		**-u-*type *synъ 'son'*
		rabъ 'slave'	*rodъ* 'birth'	
NOM.SG.	*žen-a* (< **-ā*)	*rab-ъ* (~ PIE **-os*)[a]	*rod-ъ*	*syn-ъ*
ACC.SG.	*žen-ǫ* (< **-ām*)	*rab-ъ* (~ PIE **-om*)	*rod-ъ*	*syn-ъ*
GEN.SG.	*žen-y*	*rab-a*	*rod-a*	*syn-u*
...	...	...	...	...

[a]The genesis of the actually attested endings of the nom.sg. and acc.sg. forms of the **o*-declension represents an intricate problem on its own; for a discussion and survey of the relevant literature, see, in particular, Kortlandt 1983: 181ff.; Vermeer 1991; Orr 2000: 96–113. While the development **-om* > *-ŭ* (*-ъ*) in the accusative must be regular, the expected reflex of the nominative ending **-os* should be ***-o* (preserved in some isolated forms, such as the Russian proper name *Sadk-o*). Later this ending must have been replaced by *-ŭ* (*-ъ*), in analogy with the **u*-declension, where the nominative–accusative merger was phonetically regular (**-us*, **-um* > *-ŭ* (*-ъ*)). For a convincing substantiation of this scenario, see Vermeer 1991.

Table 28.5. Nominative, accusative, and genitive plural in Old Russian

	**-ā-*type	**-o-*type	(by the 15th cent.)
NOM.PL.	*žen-y* 'wife'	*rab-i* 'slave'	→ *rab-y*
ACC.PL.	*žen-y*	*rab-y*	
GEN.PL.	*žen-ø*	rab-ъ	
...	...	...	

distinct nominative and accusative forms fell together. The only important exception is the declension type in *-a-* (< PIE **-ā-*), where the distinction was preserved. Thus, for instance, in Old Church Slavonic (around the ninth century AD), within the singular paradigm, the nominative and accusative forms are only distinguished for the *-a*-type (Table 28.4).

The Old Russian paradigm (attested since the eleventh century AD) is very similar. In the plural, after the replacement of the old nominative ending *-i* by *-y* of the accusative (approximately by the fifteenth century), the syncretism of the nominative and accusative cases has proceeded even further, since the two forms have merged not only for masculine but also for feminine nouns in *-a-* (Table 28.5).

While for the neuter nouns the syncretism of nominative and accusative was normal and even obligatory already in Proto-Indo-European, for other nouns such an innovation apparently could not be tolerated. For the resolution of the case conflict, the form of the genitive case has been taken, replacing the old accusative

Table 28.6. The category of animacy in Modern Russian

	fem.	masc.anim.	masc.inanim.
NOM.SG.	*žen-a*	*rab-ø*	*rod-ø*
ACC.SG.	*žen-u* (< *-*ǫ*)	*rab-a*	*rod-ø*
GEN.SG.	*žen-y*	*rab-a*	*rod-a*
…	…	…	…
NOM.PL.	*žen-y*	*rab-y*	*rod-y*
ACC.PL.	*žen-ø*	*rab-ov*	*rod-y*
GEN.PL.	*žen-ø*	*rab-ov*	*rod-ov*
…	…	…	…

formally identical with the nominative.[16] This innovation was limited to animate nouns (cf. the shaded boxes in Table 28.6), which thus have become opposed to inanimates. As a result, the Slavic languages have acquired the category of animacy.

Note that in the singular of the *-a*-type (but not in the plural), the old form of accusative has been preserved. Apparently, because of the lack of merger, it was not necessary to replace it with the genitive.

Here evidence from texts is particularly valuable: we can observe a continuous growth of the tendency and the expansion of the genitive. For instance, Old Russian attests the introduction of the gen.pl. forms in the function of the acc.pl. in the twelfth century for masculine human nouns; some Slavic languages, such as Polish, have stopped at that stage. The feminine humans and non-human animates introduce gen.pl. for acc.pl. only in the thirteenth century.

28.3 The main evolutionary types of languages

The analysis of possible changes in case systems presented in sections 28.1–2 can serve as a basis for classification of languages according to the general tendency

[16] In fact, the use of the genitive for encoding the direct object (exemplifying a non-canonical object marking) is quite an old feature of this group of Indo-European languages, undoubtedly going back to the Common Slavic period. Already in the oldest attested Slavic language, Old Church Slavonic (the earliest texts are dated to the ninth century AD), genitive objects were common in some types of constructions, in particular under negation and with some non-canonical transitives, such as *xotěti* 'wish'.

which determines the evolution of case systems. We can distinguish between the following three main evolutionary types of languages:

I. **Case-increasing languages** (that is, languages which undergo case-increasing) include, for instance, Uralic, New Indo-Aryan, Tocharian;
II. **Case-reducing languages:**[17] Germanic, Italic/Romance and Celtic, Albanian, Greek;
III. **Case-stable languages**: Armenian, Slavic, Baltic (Lithuanian), Turkic.

The border between case-reducing and case-stable languages cannot always be drawn with accuracy. Thus, as mentioned in the discussion of the history of the Lithuanian case system (section 28.1.1.2), the ablative and (old) locative have been lost by the time of the oldest attested texts. This feature, indicating per se the case-reducing type, was, in a sense, compensated by the rise of new locatives, and, accordingly, there are good reasons to consider (Old) Lithuanian an example of the case-stable type.

Of particular interest is the distinction between case-reducing languages, on the one hand and the case-stable/case-increasing type, on the other, which can be found within one single language family or even within a smaller genetic unit, a group of very closely related languages. Most importantly, the case-stable (or case-increasing) type by no means correlates with phonological conservatism. In other words, in the cases where two genetically related languages A and B undergo some crucial phonetic changes which, eventually, should result in the erosion of inflection, this does not yet guarantee that these processes will be equally fatal for their case systems. In reality, one of the languages may indeed be affected by heavy losses of cases or even the total collapse of case system, whereas in another one, the case system may remain essentially intact.

Thus, both Romance and Slavic languages have been subject to the erosion of case inflection, which has resulted in the merger of some case endings (see sections 30.1 and 28.2.2). However, in contrast to the Romance languages, Slavic shows a greater degree of morphological conservatism, using several compensating techniques (Harris and Campbell's 'preservative reanalysis'). As a result of one such compensating process, Slavic has developed the category of animacy, which has helped to save the nominative–accusative contrast. The only outsider within the Slavic language group is Bulgarian/Macedonian, which has lost the majority of the case distinctions, only preserving traces of the nominative ('direct'), genitive ('oblique'), and vocative forms (see e.g. Bubeník and Hewson 2006: 195–8).

Another instructive example is provided by the history of the Indo-Aryan group. Like many other Indo-European languages, the Indo-Aryan languages had lost most

[17] The loss of cases and decay of case systems (case-reducing) is dealt with in Chapter 30 and therefore is not discussed here.

978–0–19–920647–6 29-Spencer-ch28 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 456 of 457 May 16, 2008 18:59

case distinctions by the end of the Middle Indic period. Nevertheless, in contrast to the Western Indo-European languages, in the New Indo-Aryan period we observe a strong tendency to develop new cases from postpositions, which has resulted in the restoration of case systems nearly up to the previous (Old Indo-Aryan) size in such languages as Sinhala.

It seems that one of the factors that determine the evolutionary type of a language is the areal rather than genetic relationship. Thus, Baltic and Slavic language groups form a remarkable exception within the Indo-European family, being most conservative as far as the case systems are concerned. Old Lithuanian has even extended its case system, developing four new locatives (still preserved in the most archaic dialects); likewise, Russian has expanded the original (Common Slavic) case system (from seven to nine units). Both phenomena are likely to be due to Finno-Ugric influence (see Mathiassen 1996: 38).

By contrast, the collapse of the Bulgarian case system seems to represent one of the features of the Balkan linguistic area: nearly all languages belonging to this 'Sprachbund' have considerably reduced their case systems ending up with two to four cases (cf. Albanian, Greek, and Romanian).

Another instructive example is provided by the history of the Indo-Aryan case system. Here, the emergence of new cases on the basis of the multilayer case technique may be due to the influence of the adjacent Dravidian languages. Similar mechanisms may be responsible for the agglutinative restructuring of the Tocharian case system, probably due to the influence of a language or languages of the agglutinating (Turkic?) type. Finally, the evolution of the Ossetic case paradigm may be due to the influence of Caucasian (Daghestan?) languages with their rich case systems (see e.g. Cheung forthcoming).

It is interesting to note that the shared characteristics of two or more genetically related or geographically adjacent languages are usually not limited to the evolutionary type of language, which determines the global tendency in the development of case system (case-increasing, case-reducing, case-stable), but extend to some 'minor' features.

Thus, the affinity of Finnish, (Old) Lithuanian, and Russian is not limited to the case-increasing type (probably induced by Finno-Ugric). In addition, Baltic (Old Lithuanian) seems to have borrowed from Finnish the very *mechanism* of case expansion (multilayer case-marking, rather uncommon for the Indo-European linguistic type in general) and extended the same semantic area (locative) as the adjacent Finnish. The functions of the two new Russian cases, second locative and second genitive (partitive) also seem to indicate the Finno-Ugric influence. It is a commonplace in the Russian historical grammar that the rise of the partitive case is due to the influence of the Finno-Ugric case systems (such as that of Finnish, which has a partitive case). The rich system of Finnish locative cases may also have been an indirect reason for developing a new (second) locative

case, distinct from the Old Russian locative, which bears too many non-locative functions.

Acknowledgements

I am much indebted to P. Arkadiev, W. Boeder, V. Bubeník, R. Derksen, F. H. H. Kortlandt, A. Lubotsky, A. Malchukov, H. Martirosyan, A. Spencer, and M. de Vaan for critical remarks and valuable comments on earlier drafts of this paper.

CHAPTER 29

GRAMMATICALIZATION OF CASES

BERND HEINE

29.1 The role of grammaticalization in structuring case marking

The concern of this chapter is not with case systems nor with the syntactic structures associated with case expressions; the chapter deals simply with the fate of individual case markers. Space does not allow for appropriate exemplification of the grammaticalizations to be discussed, for which see the works cited in the references. We will say that case markers are grammatical forms, typically nominal suffixes (or clitics), whose main function is to assign a case property to the noun or noun phrase they govern. Our interest is not restricted to the core case functions A (subject of transitive clauses), O (object of transitive clauses) and S (subject of intransitive clauses); rather, we will in the same way be concerned with peripheral case functions, in particular with ablative, allative, dative, benefactive, comitative, instrumental, locative, possessive (genitive), and purposive functions.

While we will be looking primarily at inflectional case expressions, which in the majority of instances are suffixes, we will not attempt to trace any rigid boundary between affixes and adpositions (prepositions or postpositions), that is, free forms that serve the expression of a wider range of functional relations. Where this boundary is to be located is an issue that is notoriously controversial in both typological works and grammatical descriptions; what is described by one author as an affix is described by another author as a clitic or an adposition, and what in one language

is expressed by case suffixes may correspond to adpositions in another language or dialect of the same language.

In many, though not all, Indo-European languages, case affixes are old, they can be traced back to the earliest phases of development within this family and remain etymologically opaque. But there are other languages where their development can be reconstructed, and in this chapter we will be concerned mainly with the latter. This development is – with very few exceptions – unidirectional, being in accordance with the following principles of grammaticalization (see Heine and Kuteva 2002; 2005):

(a) **Extension**: When developing into case markers, lexical items are extended in their use to a larger range of (complement) nouns and their meaning becomes more general, and there may arise a novel meaning that is suggested by the new contexts.
(b) **Desemanticization**: Case markers lose their lexical meaning and assume a schematic case function. For example, the Balto-Finnic noun **kansa* 'people', 'society', 'comrade' lost its lexical meanings when it was grammaticalized to a comitative–instrumental case marker *-ga/-ka* 'with' in Estonian (Stolz 2001c: 599–600).
(c) **Decategorialization**: On the way from lexical item to case marker, the items concerned (i) lose most of the morphosyntactic properties characteristic of the lexical item, (ii) they tend to be reduced to invariable clitics or affixes restricted in their use to the position next to a noun or noun phrase, and (iii) they change from a morphological paradigm having many members to one having only a small number of members. Being grammaticalized to a case postposition or suffix in the modern Balto-Finnic languages, the Balto-Finnic noun **kansa* lost the categorical properties of a noun, including the ability to occur on its own, turning into an appendage *-ga/-ka* of other nouns in Estonian.
(d) **Erosion**: The case markers tend to lose phonetic substance, including the ability to carry stress, and/or they may lose their phonological autonomy, adapting phonetically to their host noun; erosion may be due to what is traditionally described in terms of *auslaut* phenomena. The Balto-Finnic example mentioned above illustrates this process, leading from a full form **kansa* to a phonetically reduced form *-ga/-ka* in Estonian, and in the Salis dialect of Livonian it was further reduced to *-k* (Kahr 1976: 117). The Persian noun root *rādi* 'reason, goal' first developed into a postposition 'by reason of, concerning' and, turning into a case marker for indirect objects and specific direct objects in modern Persian, was reduced to *rā* (Blake 1994: 167). Erosion reaches its endpoint when the case marker is lost.

As a rule, most or all of these principles are involved in the development of case markers. For example, when the ablative preposition *az* 'from' of the Iranian language Bartang was grammaticalized to an accusative case prefix *a-* (Dixon 1994:

203), this involved desemanticization of the ablative meaning, decategorialization of morphosyntactic properties of a free-standing form as an affix, and erosion from *az* to *a*. There tend to be some correlations between these principles. Adpositions are less desemanticized and decategorialized than case affixes (see (1) below), hence they also are likely to show less erosion. For example, Hungarian postpositions are (with two exceptions) bi-syllabic, while all case affixes are mono-syllabic (Kahr 1976: 119).

29.2 The rise of new case markers

Ultimately, the main sources of case markers are nominal and verbal forms, to a lesser extent also adverbs. Overall, there is the following general directionality in the grammaticalization of case markers (Heine et al. 1991; Blake 1994: 163):

(1) noun, verb (> adverb) > adposition > case affix > loss

The transition from one stage to another in (1) is gradual; for example, what is an adposition in one language can be a case affix in another closely related language or dialect. The Finnic nominal phrase **kerða-lla* 'at a turn, at time of' has given rise to the postposition *keralla* 'with' in Finnish (e.g. *koira-n keralla* [dog-GEN with] 'with the dog') and via erosion to the comitative suffix *-ke* in the related language Karelian (e.g. *velle-ŋ-ke* [brother-GEN-with] 'with the brother'; Blake 1994: 167).

Loss of a case marker does not necessarily mean that the marker disappears entirely; not uncommonly, the marker merges with another case marker that is newly created on the basis of (1). While in Latin, the locative case was lost, syncretizing with the ablative, in the closely related Osco-Umbrian, the non-distinctive locative suffix **-ei* did not disappear but merged with the postposition *en* 'in, upon' (cognate to the Latin preposition *in*; Kahr 1976: 115).

29.2.1 Noun > case marker

Nouns (or noun phrases) provide one of the main sources for case markers; Table 29.1 provides common nouns that may grammaticalize into case markers.[1] But there is a wealth of additional nouns that may develop into case markers, as Table 29.2 shows with examples from the Indo-Aryan language Bengali. This

[1] Only concrete lexical items are listed in Table 29.1, that is, we are ignoring relational nouns such as 'top', 'bottom', 'inside', etc. We are also ignoring the fact that in many languages it is not a noun on its own but rather a noun which already contains a case marker that is grammaticalized to a new case marker (see Kahr 1975 for discussion).

978–0–19–920647–6 30-Spencer-ch29 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 461 of 469 May 16, 2008 18:59

Table 29.1. Some common de-nominal sources of case markers

Nominal meaning	Case function	Typical English gloss of case function
'back', 'buttock'	locative	'behind'
'flank', 'side'	locative	'beside', 'next to'
'ground', 'reason'	cause	'because'
'head', 'eye', 'front', 'breast'	locative	'in front of'
'hand', 'comrade'	comitative	'(together) with'
'house', 'home'	locative, possessive	'at', 'of'
'place', 'area'	locative	'at', 'around'
'soil', 'ground'	locative	'below, underneath'
'stomach,' 'guts'	inessive	'in'
'top', 'head', 'sky'	superessive	'on'

Table 29.2. Presumable Sanskrit nominal sources of Bengali case markers (based on Kahr 1976: 125–6)

Sanskrit noun	Bengali case suffix
antaḥ or *antar* 'inside'	*-(ē)tē* locative
kakṣa 'hiding place, armpit'	*-kē* dative
kārya 'to be done; work, task'	*-er* genitive
madhya 'middle'	*-mi* locative (dialectal)

process, which normally involves an intermediate stage where the noun (phrase) serves as an adposition, or an adverb, entails most of all desemanticization, whereby the lexical semantics of the noun is lost or, more precisely, is reduced to expressing a case property. Decategorialization has the effect that the noun, due to the fact that it is attached to another noun, loses its combinatorial potential and its internal morphological complexity and tends to be reduced to an invariable grammatical form serving some specific syntactic function.

But there is also a more general process whereby nouns develop into adpositions and some of the latter may turn into case clitics and case inflections (see below). Most cases that have been documented involve nouns that give rise to locative adpositions and finally to locative case inflections, although the use of the latter may be, and frequently is, extended to also denote temporal, causal, and other case relations. But it is not only the spatial domain that is involved. The example mentioned above illustrates an alternative pathway of grammaticalization: The Balto-Finnic noun **kansa* 'people', 'society', 'comrade' developed into the comitative postposition *kanssa* in Finnish and *kaas* ('together with', 'in the company of') in Estonian, and eventually it turned into a comitative–instrumental marker *-ga/-ka*

in Estonian. In a similar fashion, the comitative marker *-(gu)in* of the fellow Finnic language Sami appears to be etymologically derived from the Sami noun *guoibmi* 'comrade, fellow, mate' (Stolz 2001c: 599–600; see Heine and Kuteva 2002: 91–2 for additional evidence for this grammaticalization). A similar development appears to have happened in Basque, where the noun *kide* 'companion', 'fellow', 'mate', applied to both people and things, appears to be the source of the comitative case suffix *-ekin* (Heine and Kuteva 2002: 91–2).

Rather than the meaning of the noun, it may be inflections on the noun that determine the function of the resulting case affix. The Hungarian case suffixes *-ben/-ban* 'in' (inessive) and *-ból/-bo͡l* 'away from' (elative) are both historically derived from a relational locative noun *bél* meaning 'interior'.[2] The difference in case functions is due to the fact that the final segments *n* and *l* are themselves relics of the case suffixes *-n* locative and *-öl* ablative on the locative noun (Lehmann 1982: 84–5; Hopper and Traugott 1993: 107–8).

The endpoint of the grammaticalization path from noun to case marker may be that the latter ends up as a meaningless appendage of nouns or some other elements. In the Finnic language Komi-Permiak, the erstwhile noun *vyv* 'top' was grammaticalized and survives only as the first formative *-vv-* in five case suffixes (Blake 1994: 167).

29.2.2 Verb > case marker

Together with nouns, verbs provide the most common source of case markers (Kahr 1975; 1976). A list of common grammaticalizations is found in Table 29.3; note that this list is far from exhaustive, there is a vast range of verbs in the languages of the world that have given rise to case markers.

In the Uto-Aztecan language Ute, all locative case markers originate from historically still traceable precursor verbs which have turned into noun suffixes, no longer carrying any discernible residue of verbal properties (Givón 2006: 24); Table 29.4 is a selection of these case markers and their lexical sources.

29.2.3 Adposition > case affix

As we saw in (1), adpositions constitute an intermediate stage on the way from lexical item to case marker, and the vast majority of case affixes go back to adpositions. For example, for all Hungarian case affixes that are historically documented, the source is a postposition (Kahr 1976: 121); thus, Lehmann (1982: 85) observes that in its preliterary period before AD 1200, Hungarian had postpositions; but

[2] The original meaning was presumably 'innards, guts' (Blake 1994: 167).

978–0–19–920647–6 30-Spencer-ch29 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 463 of 469 May 16, 2008 18:59

Table 29.3. Some common de-verbal sources of case markers

Verbal meaning	Case function	Typical English gloss of case function
'come from', 'leave'	ablative	'from'
'be at'	locative	'at'
'go to', 'come to', 'reach', 'arrive at'	allative	'to'
'give'	benefactive, dative	'for', 'to'
'meet', 'join', 'follow'	comitative, instrumental	'with'
'take'	patient, object case	–
'pass'	perlative, pergressive, path	'through', 'along'

Table 29.4. Some verbal sources of locative case markers in Ute (Givón 2006: 24)

Verbal source	Case suffix
-cawi 'come to'	*-caw* 'toward'
-kwa 'go to' (defective verb)	*-kwa* 'to'
-naagha 'enter'	*-naagh* 'in'
-rukwa 'descend'	*-ruk* 'under'
-tarugwa 'climb'	*-tarux* 'on (top)'

as from the beginning of the literary tradition, the postpositions appear as case suffixes.

The development from adposition to case affix typically involves the following principles of grammaticalization. Extension has the effect that the use of an adposition is extended from occurring with a limited set of nouns to a larger (or unlimited) set of nouns. Via desemanticization, the adposition tends to lose the specific semantics it may have had and is reduced to some schematic case function, and decategorialization means that the adposition loses any morphosyntactic autonomy it may have had, turning into an appendage of nouns (or noun phrases). In addition, there also tends to be erosion, whereby the adposition loses in phonetic substance, including the ability to carry stress. The result of erosion is that case affixes are on average shorter and are more likely to assimilate phonetically to the noun to which they are attached.

It is most of all adpositions denoting location, source, or destination that develop into case markers. In the Iranian Pamir language Bartang, the ablative preposition *az* has become grammaticalized as an accusative case prefix *a-* (Dixon 1994: 203),

and in the Penutian language Nez Perce, the ergative case marker *-nim* appears to be historically derived from an adpositional particle meaning 'hither' (Blake 1994: 169).

29.2.4 Adverb > case marker

This pathway of grammaticalization is commonly encountered in Indo-European languages, leading from adverbs to adpositions, and in some cases also to case markers. For example, in Hittite, a number of locative adverbs developed into case-sensitive postpositions, like the Old Hittite adverb *anda* 'into, inside' which gave rise to the postposition *anda* 'into' (Bubeník n.d.).

29.2.5 Possessive cases

Possessive (genitive) case markers tend to occupy cross-linguistically a marginal position in case paradigms, being part of the syntax of noun phrases rather than of clauses. Many of these case markers, like the genitive case inflections of Indo-European languages, are etymologically opaque, that is, their genesis is beyond the scope of the methodology of historical linguistics. But there are possessive case markers and the possessive constructions associated with them for which there exists sufficient diachronic evidence to allow for generalizations on their evolution. The main conceptual schemas serving as sources are listed in Table 29.5 (Heine 1997b; see also Godel 1975: 106; Heine and Kuteva 2002: 34–5; Kortlandt 2003: 47).

Three of the schemas listed there (location, source, and goal) can be described in a broad sense as being spatial in nature, whereby the possessor is conceptualized as a spatially described participant. In European languages, the schema predominantly recruited is Source, where the possessor is presented by means of an ablative preposition ('(away) from', 'out of'), e.g., English *of* (< *off*), German *von*, Dutch

Table 29.5. The main source schemas used cross-linguistically for the expression of attributive possession and of possessive cases (adapted from Heine 1997b: 144)

Formula	Label of event schema
Y *at* X	Location
Y *from* X	Source
Y *for/to* X	Goal
X *with* Y	Companion
(*As for*) X, X's Y	Topic

van, Frisian *fan*, Catalan *de*, Macedonian *od*, Upper Sorbian *wot*, etc. Thus, there was a historical development whereby a locative prepositional construction of the form [*from* NP] developed into a possessive/genitive construction [GEN possessor]. For example, the Latin prepositional construction [*de* X 'from X'] is the historical source of attributive possessive constructions to be found in the modern Romance languages. In a number of other European languages, the location schema was used to create possessive constructions, where a static locative preposition 'at' was grammaticalized to a possessive case marker, such as Faroese *hjá* 'at', Scottish Gaelic *aig* 'at', Irish *ag* 'at', or Albanian *prej* 'at'.

Another source for possessive case markers can be seen in nouns meaning 'property', 'part', or 'thing'. The process underlying this development is based on the reinterpretation of a structure [X *property (of)* Y] as [X *of* Y], whereby the noun is desemanticized and decategorialized to a functional marker expressing a syntactic relation. In the Aztecan language Pipil of El Salvador, the relational noun *-pal* 'possession' has turned into a preposition *pal*, and a possessive case marker (Harris and Campbell 1995: 126–7), and the Arabic noun *bita:* ʕ 'property' has provided the source for the genitive marker *ta* in the Arabic-based creole Nubi (Heine and Kuteva 2002: 245–6).

29.2.6 Discussion

The form of case markers may still bear witness of their lexical origin. First, case markers derived from nouns tend to retain some relics of a possessive/genitive morphology, while verb-derived case markers may contain some nominalizing marking. Thus, in English adpositional case markers the genitive/possessive particle *of* (*in front of*, *on top of*, etc.) reflects their nominal, and the present participial ending *-ing* (*regarding*, *following*, *concerning*, etc.) their verbal origin. In a similar fashion, the Finnish genitive suffix *−n* on the noun reflects the nominal origin of the adessive *kohdalla* 'at the place', as in *talo-n kohdalla* (house-GEN place. AD) 'at the house' (Blake 1994: 167). Second, the position of case markers is likely to reflect the syntactic phrase structure from which they are derived. For example, in the Niger–Congo language Ewe of West Africa, noun-derived case markers follow while verb-derived case markers precede the noun they govern; thus, in the phrase *le xɔ-á megbé* (at house-DEF behind) 'behind the house', *le* is a preposition derived from the verb *le* 'be at' while *megbé* is a postposition derived from the noun *megbé* 'back', and the postposition cannot be used used without a valency-sensitive element such as the preposition *le*. This ordering reflects the respective historical source structures: while prepositions can be reconstructed back to a structure [verb–N], postpositions are derived from a construction [genitive modifier N–noun] (Heine et al. 1991). And third, in a number of languages, verb-derived case markers, but not noun-derived

case markers have valency properties; for example, the noun-derived postposition *megbé* of Ewe requires the verb-derived preposition *le* as a valency marker.

The pathways sketched in section 29.2 are not the only ones through which case markers may arise; there is a range of additional sources that need to be distinguished. One not uncommon source domain is provided by pragmatically marked constructions which may undergo syntacticization, whereby functions of information structure, such as topic and focus, are grammaticalized to case markers, in particular subject/nominative markers. For example, in the W2 dialect of the North Khoisan language !Xun, the particle *má*, marking sentence topics, has been grammaticalized to a subject marker in specific contexts. Another source of case markers is provided by definite articles (Harris and Campbell 1995: 341); in the West Nilotic languages Anywa, Päri, and Jur-Luwo in northeastern Africa an earlier definite marker has developed on the one hand into a marked nominative and on the other hand into an ergative case marker (König 2006; 2008a).

29.3 From one case function to another

Existing case markers commonly acquire new case functions via the grammaticalization principle of extension (see above), whereby the use of a case marker is extended from one syntactic context to another, thereby giving rise to case 'polysemy'. This section is concerned with major directions in the development of case forms. Table 29.6 lists the most common kinds of case extensions. To illustrate these grammaticalizations would require more space than is available; the reader is referred to the following works in particular for exemplification: Blake (1994), Dixon (1994), Heine et al. (1991), Harris and Campbell (1995), Heine and Kuteva (2002), König (2008a), Narrog (this volume, Chapter 40), Narrog and Ito (2007).

The grammaticalizations listed in Table 29.6 allow for the following generalizations: First, case markers and their functions can be defined with reference to their relative degree of grammaticalization. Thus, the case functions in the left column are less grammaticalized than the one in the right column. For example, a common extension process pointed out by Dixon (1994: 202) is from A to S, and – less commonly – from O to S. This would seem to suggest that, of all the three case functions, S is the most strongly grammaticalized function. Second, a given case function may derive historically from more than one other case function; for example, partitives commonly derive either from ablatives or possessives (Harris and Campbell 1995: 339–40). And third, it is possible on the basis of their relative

degree of grammaticalization to arrange case functions in the form of chains of grammaticalization (see (2) below).

Ergative case markers have a number of different sources. A not uncommon source is provided by agent case markers in passives that develop into ergative constructions, another source is provided by instrumental markers or genitive case markers in possessive constructions (Dixon 1994: 190–1, 204). When a nominal structure is reinterpreted as a clausal structure, it may happen that a genitive case marker, used to introduce a nominal modifier, is reinterpreted as an ergative, agent, or even a patient case marker.

While they are frequently derived from other case functions, ergatives may themselves develop further and assume other case functions. In the Caucasian (Kartvelian) language Mingrelian, the original ergative case marker was extended to cover both S and A in the aorist, resulting in a nominative–accusative system (Dixon 1994: 202).

Extension from one case function to another is the primary – though not the only – source for case polysemy. Almost all languages of western Europe exhibit a polysemy involving comitative ('together with') and instrumental ('by means of') case functions (see Stolz 1996a; 1996b; 2001c), and nearly two-thirds of the documented languages of Australia have a suffix *-gu* expressing allative, dative, and/or genitive functions on nouns and purposive and future functions on verbs (Dixon 2002: 166–7). That such polysemies are due to grammaticalization has been argued for by e.g. Heine and Kuteva (2006): new comitative–instrumental polysemies arise via the grammaticalization from comitative to instrumental case marker.

On the basis of their relative degree of grammaticalization, it is possible to arrange case markers and their functions in the form of semantic maps (see Narrog and Ito 2007) or chains of grammaticalization (Heine et al. 1991). Thus, the grammaticalizations summarized in Table 29.6 suggest the following fairly common sequences of grammaticalization:

Table 29.6. Common patterns of extension from one case function to another

From	To
A	S
ablative	cause, possessive, partitive, instrumental
allative	benefactive, dative, accusative/0, purposive
benefactive	purposive
comitative	instrumental, ergative, manner, possessive
dative	accusative/0
instrumental	ergative, manner
locative	comitative, agent, ergative, instrumental

978–0–19–920647–6 30-Spencer-ch29 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 468 of 469 May 16, 2008 18:59

(2) Some common grammaticalization chains of case functions
   a. allative > benefactive > purpose
   b. allative > dative > accusative/O
   c. locative > comitative > instrumental > manner

29.4 From case marker to non-case marker

Case inflections are commonly derived from other categories, as we saw in section 29.2, and they usually are fairly strongly grammaticalized forms. But they themselves may be deployed for further pragmatic and syntactic functions, giving rise to other functional categories. There are in particular the following directions in which case markers may further grammaticalize: clause subordination, modality, noun phrase conjoining, and tense marking.

Perhaps the most salient pathway away from case marking consists in the grammaticalization pathway from case marker to clause subordinator. This development is usually based on the extension principle, whereby the use of case markers is extended from nouns to nominalized (infinitival, gerundival, or participial) verbs, and finally to subordinate clauses. For example, the ergative/instrumental marker *-na* of the Tibeto-Burman language Newari, described as a postposition of nouns (3a), is hypothesized to have given rise to the temporal clause subordinator *-na* suffixed to verbs (3b).

(3) Dolakhari Newari (Tibeto-Burman; Genetti 1991: 227)
   a. *cotan- **na** pol- ju.*
      spoon- INS strike- 3SG.PAST
      'He eats with a spoon.'
   b. *chẽ- ku yer- **na** wā ām- e naku*
      house- LOC come- WHEN EMPH he- GEN cheek
   c. *moŋ- an coŋ- gu.*
      swell- PART stay- 3SG.PAST.HAB
      'When he came to the house his cheek was swollen.'

The result is that in quite a number of case languages, at least some types of subordinate clauses are marked by what are historically noun case markers. For example, the dative case suffixes *-k^e* of the Kuliak language Ik of Uganda and *-rò* of the Saharan language Kanuri of Nigeria have been grammaticalized to fairly general markers of clause subordination (Heine 1990), and Dixon (1994: 192) observes that in the Carib languages of South America, subordinate clauses generally have

978–0–19–920647–6 30-Spencer-ch29 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 469 of 469 May 16, 2008 18:59

the status of nominalizations and show an ergative pattern. The second direction is immediately connected with the first one: Subordinate clauses tend to express modally marked meanings, referred to with terms such as 'subjunctive' or 'irrealis'. Now, when a case marker is grammaticalized to a clause subordinator it may acquire the modal meanings typically associated with subjunctive moods.

The third common pathway of grammaticalization leads from comitative case markers ('together with') to markers conjoining noun phrases ('and'), whereby a phrase [X *(together) with* Y] is grammaticalized to a structure [X *and* Y] (Stassen 2000). And the third directionality leads from case to tense: purposive case markers are not uncommonly extended from nouns to verbal constitutents and can develop further into future tense markers (Bybee et al. 1994; Dixon 2002: 166).

29.5 Case behaviour in language contact

That the behaviour of case categories is sensitive to language contact is nowadays fairly uncontroversial (see also Chapter 32, this volume). Language contact may on the one hand have the effect that existing case categories are given up or new case categories are created on the model of some other language; see Kahr (1976: 141–3) for examples. On the other hand, contact may lead to the further grammaticalization of existing case markers. The following pathways have been observed in situations of language contact, and these pathways apply irrespective of whether inflectional or any other forms of case marking are involved (see Heine and Kuteva 2005; 2006 for examples):

(4) Contact-induced grammaticalization of case markers
   a. comitative > instrumental
   b. allative > dative marker
   c. dative > accusative/O marker
   d. peripheral > core participant marker

As has been shown in recent studies (Heine and Kuteva 2005; 2006), contact-induced changes in case functions follow much the same principles of grammaticalization as changes not involving language contact.

CHAPTER 30

CASE IN DECLINE

JÓHANNA BARÐDAL
LEONID KULIKOV

This chapter focuses on the main aspects of the reduction/loss of case and the decay of case marking systems. The general mechanisms which lead to the merger of case and case syncretism and, eventually, to the loss of (some) cases include: (i) phonetic processes which result in the loss of the difference between two or more case forms, i.e. erosion of case inflection, and, thus, in case syncretism; (ii) overlapping of syntactic and semantic functions and/or uses of individual cases, i.e. syntactic and semantic affinity of some cases; (iii) semantic or functional overlapping of whole argument structures; and (iv) a variety of analogical developments and paradigmatic levelling (cf. Kulikov 2006).

Often these mechanisms work together so that several factors create favourable (albeit not always sufficient) conditions for the case mergers. The phenomenon of case syncretism can be best illustrated with examples from the history of the Indo-European languages which attest nearly all possible types of case mergers within the original eight-case Proto-Indo-European case system: genitive–ablative (Slavic, Greek), nominative–accusative (Balkan: Romanian, Albanian), dative–locative (Greek), ablative–instrumental–locative (Latin), dative–ablative–instrumental–locative (Celtic, Germanic), etc. (cf. Luraghi 1987 and Chapter 14, this volume). The ultimate case syncretism is typically preceded by a period of variation and alternation between case forms or argument structures, with the source forms being interchangeably employed in some usages with only some minor functional distinctions (see Kulikov, to appear). The interplay between phonetic erosion and the semantic/functional overlap of case forms and argument structure

978–0–19–920647–6 31-Spencer-ch30 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 471 of 478 May 16, 2008 18:59

constructions can be demonstrated with examples from several Indo-European language groups, as these provide rich evidence for various scenarios of the decay and collapse of case systems.

30.1 Phonetic erosion of case inflection

In the simplest and most trivial cases, the (partial) merger of case morphemes and, eventually, the decay of case systems is due to certain phonological changes, foremost, to the erosion of inflection in word-final position (in languages with case suffixes) or, much more rarely, in word-initial position (in languages with case prefixes). Such a development may result in case syncretism, where case distinctions are erased in their entirety.

The evolution of the Arabic nominal inflection provides an instructive example. In the post-classical period, Arabic undergoes a strong reduction of case endings, resulting in the loss of the original three-case system. Phonologically, these processes essentially amount to the weakening, merger, and the subsequent loss of final vowels (in particular, Nom.Sg. *-u*, Gen.Sg. *-i* and Acc.Sg. *-a*). Middle Arabic of the Southern Palestinian Christian texts of the eighth–tenth centuries AD still exhibits vestiges of case distinctions, although the oppositions of the classical language appear severely deteriorated. One case variation found in this period is that between the genitive, accusative, and nominative on nominal forms preceded by prepositions (where case endings were preserved longer than in many other contexts), as illustrated in (1):

(1) Southern Palestinian Christian Middle Arabic (Gruber-Miller 1990: 244f.)

a. *w-l-ʔb-**ii**-h*
and-to-father-GEN-his
'… and to his father'

b. *mʕ ʔb-**aa**-hmaa*
with father-ACC-their
'… with their father'

c. *y-tklm ʕlaa ʔx-**uu**-h*
3MASC.SG.IMPF-speak against brother-NOM-his
'He speaks against his brother …'

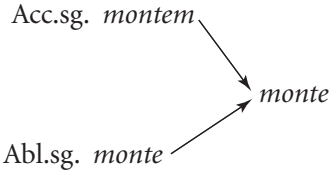
Another example of the total collapse of a case system, primarily due to phonetic developments in word-final position and erosion of case endings, can be seen in the history of the Proto-Romance (i.e. Latin) case system in the daughter languages. Latin attests the very beginning of the decay of the original Proto-Indo-European

978–0–19–920647–6 31-Spencer-ch30 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 472 of 478 May 16, 2008 18:59

case system (see Section 30.2 below), which has affected all Romance languages. The daughter languages, i.e. Spanish, Italian, French, Romanian, etc., display the same tendency, reducing the Latin case system further, ending up with caseless systems or with two cases at maximum as in Romanian (Penny 2002; Calabrese 1998; Blake 2001: 175f.; Bubeník and Hewson 2006: Ch. 11). This can be shown with an example from Spanish where the relevant phonological changes are the following:

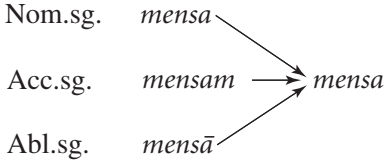
(i) Loss of final *-m* mostly causing Acc.sg. to merge with the Abl.sg.:

-m > ø :



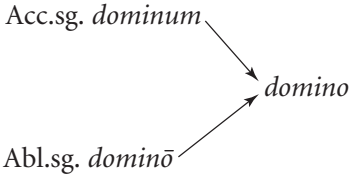
(ii) The merger of the long and short *a*, together with the loss of final *-m*, caused the merger of Nom., Acc. and Abl.sg.

-m > ø , *ā* > *a* :



(iii) The merger of *u*(*m*) and *ō* in final position caused the merger of Acc.sg. and Abl.sg.:

u(m), ō > *o* |__ # :



(iv) The merger of the front vowels in final position caused the merger of Nom.-Acc.pl. (*montēs*) with Gen.sg. (*montis*).

By the fourth–fifth centuries AD these changes had resulted in a considerable reduction of the case paradigm: a three-case system in the Eastern part of the Roman empire and two cases in most of the West, including Spain. The latter entails that the three oblique cases had merged into one common form, hence the system consisted of only nominative and accusative (oblique) case, as illustrated by the three examples in Table 30.1 (Penny 2002: 114–19).

Such two-case systems survived in French (see below) and Provençal until the twelfth–thirteenth centuries (cf. Chapter 47, this volume, for a typological analysis of two-case systems). In other areas, there was a further reduction to invariable singular and plural forms. By virtue of additional phonetic changes most of the

Table 30.1. Old Spanish case system

NOM.SG	*rosa*	*annos*	*leo*
ACC.SG	*rosa*	*anno*	*leone*
NOM.PL	*rosas*	*anni*	*leones*
ACC.PL	*rosas*	*annos*	*leones*

Table 30.2. Modern Spanish nominal paradigm

	(1)	(2)	(3)
SG.	*rosa*	*año*	*león*
PL.	*rosas*	*años*	*leones*

contrasts shown in the above table have become obliterated, surviving only for *anni* ∼ *annos*. Of course, this isolated subtype could not survive for a long time, foremost due to the levelling pressure of the morphological paradigm. Accordingly, the form *annos* has been generalized as a plural form, in analogy with plural *-s* from other words in the nominal paradigm. The resulting system of the three major paradigmatic classes that Spanish inherits from Latin is represented in Table 30.2.

30.2 Functional mergers

An example of erosion of case inflection supported by functional mergers is provided by the syncretism of three Proto-Indo-European cases, ablative, locative, and instrumental, into the Latin ablative (for details of the history of the Latin case inflection, see, in particular, Leumann et al. 1977: 405ff.). The relevant fragment of the system of case endings reconstructed for Proto-Indo-European (including the endings traditionally regarded as borrowed from the pronominal paradigm) is represented in Table 30.3. The endings which have left direct reflexes in the actually attested markers of ablative are in bold face while those which have only indirectly contributed to the attested endings are bold and underlined.

The resulting system of ablative endings, arranged by declension types, is shown in Table 30.4. Although the origins of some actually attested endings may be the subject of debate, the main details of the scenario are quite

Table 30.3. Proto-Indo-European ablative, locative, and instrumental case endings

	Singular	Plural
Ablative	*-(o)s ***-ed**	***-ios**
Locative	***-i**, *-ø*	*-su, ***-oisu**[a]
Instrumental	*-(e)H_1 *-bhi*	***-ōis**[a] ***-bhi(s)**[b]

[a] With *o*-stems.
[b] For the endings of the instrumental in the proto-language and their relations with the inflection of the dative and ablative plural, see, in particular, Kortlandt 2003: 48f.

Table 30.4. Latin ablative case endings and their Proto-Indo-European sources

Declensions	Singular	Plural
1st (*-a-*)	*-ā(d)* [analogy with *-o-*type]	*-īs* [analogy with *-o-*type]
2nd (*-o-*)	*-ō(d)* <*…o-ed*	*-īs* < *-oisu* and/or *-ōis
3rd (*-i-*, *-C-*)	*-ī(d)* [analogy with *-o-*type] *-e* < *-i*	*-bus* < *-bhos* (← *-ios + *-bhis*?)
4th (*-u-*)	*-ū(d)* [analogy with *-o-*type]	
5th (*-e-*)	*-ē(d)* [analogy with *-o-*type]	

clear.[1] This example from Latin is useful as it shows that phonetic processes may render formal distinctions between cases opaque, thus leading to the merger of some forms (as in the case of Loc. and Ins.pl.), although they do not represent the *only* driving force of case syncretism. All three source cases have left their traces in both the singular and plural paradigms at least in some of the attested Latin declensions, so phonetic processes alone could not yet result in the simple syncretism of these three cases. Hence, the final outcome is a result of a complex interplay of several mechanisms; in particular, the three source cases must be considered semantically (functionally) close enough to each other, which in turn has licensed the form of one of them to take over the functions of the other(s).

[1] The genesis of the Abl.pl. ending of the 3rd, 4th, and 5th declensions *-bus* poses some problems. It is likely to represent the Proto-Italo-Celtic *-*bhos*, which replaced the original ending *-*ios*, presumably under the influence of the instrumental ending *-*bhi* (cf. Homeric Gr. -φι; see Kortlandt 2003: 50).

30.3 Analogical developments and paradigmatic levelling: Total collapse of case systems

An instructive example of a total collapse of a case system primarily based on a number of analogical developments and paradigmatic levelling is provided by the evolution of the Old French two-case system. By the Old French period only two cases have survived (usually called subject and object cases, or 'sujet' and 'régime'), as illustrated in Table 30.5 (for details and discussion, see e.g. Plank 1979; van Reenen and Schøsler 2000).

As Table 30.5 shows, each declension type counts no more than two forms in total, distributed quite intricately across the paradigm. The system becomes even more opaque because of the loss of final *-s* before a consonant:

-s > -ø | _ *C*

Thus, for *mur-* we have two allomorphic variants, given in Table 30.6.

The factors which caused further collapse of this system include: (i) the expansion of constructions with non-canonical subject marking, viz. with the subject encoded by the object case – as, for instance, in impersonal constructions of the type *Il i a* __ (i.e. Modern French *Il y a* __ 'There is __'), which has apparently triggered case variation in the subject position (see Laubscher 1921: 51ff.); (ii) the existence of

Table 30.5. Old French case system

	masculine	feminine
SUBJ.SG	*li forz mur-s* 'strong wall'	*li forz flor-s* 'strong flower'
OBJ.SG	*le fort mur*	*la fort flor*
SUBJ.PL	*li fort mur*	*les forz flor-s*
OBJ.PL	*les forz mur-s*	*les forz flor-s*

Table 30.6. Old French case system: phonetic realizations

	__V, __ ##	__ C
SUBJ.SG	[myrs]	[my:r]
OBJ.SG	[myr]	[myr]
SUBJ.PL	[myr]	[myr]
OBJ.PL	[myrs]	[my:r]

a few (minor) inflectional types which had completely lost their case distinctions by the Old French period; (iii) the very intricate distribution of as few as two markers, *-ø* and *-s*, across the four-member paradigm, which may have rendered the system as 'conceptually too complicated' (van Reenen and Schøsler 2000: 337).

30.4 Synonymous argument structure constructions

It is a well-known fact that languages have a tendency to abate synonymous grammatical forms over time. For case and argument structure, this can take place in two ways: (i) the morphological case distinctions disappear with a consequent merging of the argument structure constructions; (ii) productive case and argument structure constructions attract new verbs and verbs from non-productive constructions, thereby gradually causing non-productive constructions to fall into disuse. Given a definition of productivity based on type frequency, semantic coherence, and an inverse correlation between the two, the productivity of case and argument structure constructions is, at least in part, derived from the size/type frequency of each case construction (cf. Barðdal forthcoming, a). Hence, the case construction lowest in type frequency is expected to disappear first, then the one next lowest in type frequency, etc., until only the productive case constructions are left in the language. This development correlates in part with changes in the verbal vocabulary, as productive argument structures attract new verbs while non-productive argument structures do not. Hence, contact situations with massive replacement of the vocabulary can speed up this development. In Germanic both developmental paths outlined above are documented. In Mainland Scandinavian and English the development has led to case merging and case loss, whereas in German and Icelandic the development has led to the disuse and disappearance of the argument structures lowest in type frequency.

Table 30.7 shows case and argument structures which can be postulated for two-place predicates in Germanic on the basis of comparative evidence and documented case marking in the history of Icelandic (Barðdal forthcoming, b).

The case and argument structure construction highest in type frequency in all the Germanic languages was without a doubt the nominative subject construction, while dative subject predicates were low in type frequency and accusative subject predicates were even less common. A comparative study of the semantics of accusative and dative subject predicates across the Germanic languages reveals that they are grossly speaking either (i) stative/inchoative experience-based predicates, or (ii) anti-causative intransitives (Barðdal 2004). There was, thus, a considerable

Table 30.7. Case constructions in earlier Germanic

NOM	ACC	DAT	GEN
NOM–ACC	ACC–NOM	DAT–NOM	GEN–NOM
NOM–DAT	ACC–ACC	DAT–GEN	GEN–PP
NOM–GEN	ACC–GEN	DAT–PP	GEN-S
NOM-PP	ACC-PP	DAT-S	
NOM-S	ACC-S		

overlap in the semantics of accusative and dative subject predicates in Germanic, also found for the nominative subject construction, which was the semantically most open construction of them all. A comparison of Nom–Acc, Nom–Dat, and Nom–Gen in Modern Icelandic also reveals that Nom–Dat and Nom–Gen are not strictly confined to any particular semantic fields, but can be regarded semantically as proper subsets of the Nom–Acc argument structure construction (cf. Barðdal forthcoming, c). This comparative evidence suggests that the case and argument structure constructions in Germanic were partly synonymous.

The genitive subject construction, which was lowest in type frequency of all the subject constructions, is not documented in Old English and Old Swedish. It thus seems that it had already disappeared before recorded history. The first documented construction to disappear in Old Swedish is the genitive object construction, i.e. the construction lowest in type frequency of all the object constructions. This took place before 1350 (cf. Delsing 1991). In English, on the other hand, genitive objects disappeared in two rounds: the genitive objects of Acc–Gen and Dat–Gen disappeared during the twelfth century while genitive objects of Nom–Gen did not disappear until the thirteenth century (Allen 1995: 217–19). This is in accordance with differences in the size of these constructions as Acc–Gen and Dat–Gen were much lower in type frequency than the Nom–Gen construction. The distinction between accusative and dative on nouns, both subjects and objects, was lost in English during the thirteenth century, after the loss of the genitive. Finally, the oblique subject construction (formerly accusative and dative subject construction) starts losing ground during the fifteenth century and only exists in fixed expressions after that (Allen 1995: ch. 6). In Swedish, moreover, the accusative subject construction (which was lower in type frequency than the dative subject construction) was lost around 1400 (Falk 1997: 14–15) and *c.*1450 the case distinctions on nouns had completely disappeared. The oblique subject construction (visible on pronouns) survived in Swedish until the sixteenth and seventeenth centuries.

In German the genitive subject construction started disappearing during the thirteenth century (Seefranz-Montag 1983: 173–5). The genitive object construction has been heavily reduced in the history of German, with only a few predicates left, and so has the dative object construction, although the dative object construction

is still higher in type frequency than the genitive object construction, with perhaps around 100 predicates in total (cf. Maling 2002). The accusative and the dative subject constructions have also been heavily reduced in German, with approximately 80–100 predicates left (cf. Barðdal 2004). In the history of German, moreover, accusative and dative subject predicates have been interchangeable, with dative subject predicates attracting more verbs from the accusative subject construction than the accusative subject construction from the dative one. In summary, the construction lowest in type-frequency i.e. the genitive subject construction, has disappeared, the remaining low type-frequency constructions, i.e. genitive and dative objects, and the accusative and dative subject constructions, have gone down in type frequency. This is because the predicates instantiating the low type-frequency constructions have either disappeared in German or occur now in the Nom(–Acc) construction.

Finally, in Icelandic, only one construction has completely disappeared, namely the Dat–Gen construction, which was instantiated by only two predicates in Old Norse-Icelandic (cf. Barðdal 2001: 197–8). Three other low type-frequency constructions are at the border of becoming extinct today, namely the Acc–Nom, Acc–Gen, and Gen–Nom constructions. These were slightly higher in type frequency in Old Norse-Icelandic than the Dat–Gen construction, and are now lowest in type frequency of all the case constructions in Modern Icelandic. The Nom–Gen construction has also been reduced in the history of Icelandic. Nom–Dat predicates in Modern Icelandic are approximately 750 (Maling 2002: 31), accusative subject predicates are *c.*200, and dative subject predicates are around 700 (Barðdal 2004). Hence, only the case and argument structure constructions lowest in type frequency in Old Norse-Icelandic have disappeared, and the ones that were already low then have decreased in type frequency. The constructions of intermediate size have maintained their status (like Nom–Dat), and the most productive Nom–Acc construction has increased its type frequency (cf. Barðdal forthcoming, b).

The loss of case and the time/onset of these changes correlate with the degree of language contact found in the individual Germanic language communities during medieval times. England was exposed to the most language contact and earliest, namely during the eleventh century. Mainland Scandinavia has been exposed to less contact, beginning in the thirteenth century. Germany has had considerably less contact and more spread out in time, while Iceland, being the most isolated of the four, has been in the least contact of them all. Clearly, rapid changes in the vocabulary favour the most productive case and argument structure constructions and disfavour the non-productive ones, causing them to fall into disuse earlier.

CHAPTER 31

THE GEOGRAPHY OF CASE

BALTHASAR BICKEL
JOHANNA NICHOLS

31.1 Introduction

Cases are of course not evenly distributed worldwide. It is generally known, for example, that cases are common in Eurasia and much less common in Africa. Modern typological research aims at capturing and understanding such continent-wide frequency differences (Nichols 1992; Bickel 2007b), and it has becomes standard practice in universals research to control for confounding factors from continent-wide linguistic areas (Dryer 1989; Cysouw 2005). A fundamental problem of linguistic geography, however, is that it is all too easy for the human eye to detect spatial patterns on a map even when they are artefacts of chance or when they arise simply because some regions have many more different people and languages than others (cf. Siberia with Cameroon; e.g. Nettle 1999).

Our approach to linguistic geography starts from biogeographical and culture-historical theories of population movements and contact patterns that define a constant set of areas as predictor variables for statistical modelling (Predictive Areality Theory: Bickel and Nichols 2006). Thus, areas are not defined linguistically, and this avoids circularity when used in linguistic surveys. Hence the present chapter does not fish for areas by visual inspection of maps, but assumes areas as hypotheses

978–0–19–920647–6 32-Spencer-ch31 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 480 of 493 May 16, 2008 18:59

and asks what, if any, aspects of case structure and case behaviour are significantly different across the hypothesized areas.

In this chapter we test a set of previously established areas against thirty-five typological variables that concern case inventories and various morphological and syntactic properties of case and are drawn from the Autotyp database (Bickel and Nichols 1996ff.), the data from selected chapters of Haspelmath et al. (2005; henceforth WALS), or, when variables and their coding were identical or near-identical in the two databases, a merged set.

We begin by describing the areas (section 31.2) and the typological variables (section 31.3) tested. After explaining our method of sampling and testing (section 31.4), section 31.5 discusses the results.

31.2 Areas

All area definitions are based on previous work (for detailed maps, see the Autotyp web site[1]):

• Africa, including the Arabian peninsula, which has well-established historical continuities with Africa. This large area is generally accepted on the basis of the strong historical connections of the entire region. As such, the African macroarea figures as a standard control in universal research. We assess the areality of Africa by comparing its profile to the rest of the world.

• Europe, delimited in the east by a line starting from the northern coast of the Black Sea, following the northern slopes of the Carpathians and then following roughly the Wisła up to the Baltic Sea and including Scandinavia but not Finland. The region is defined so as to include the major internal spreads in historical times (cf. Haspelmath 2001 for some discussion in typological perspective) and to exclude the north Eurasian steppe, but we note that the population history of the steppe has always had deep inroads into Central Europe. The areality of Europe is assessed by comparing it both to the rest of the world and to the rest of Eurasia, including Southeast Asia.

• Eurasia: all of Eurasia, except the north Asian Coast, which we assign to the Circum-Pacific region (cf. below). The areality of Eurasia has first been argued for by Jakobson 1931 on linguistic grounds, but it has robust non-linguistic support: the history of Eurasia shows frequent criss-crossing spreads throughout the area in known history, cf. the various east-to-west spreads through Central Asia (Uralic, Indo-European, Turkic, etc.), the spreads associated with the Silk Road and its

[1] www.uni-leipzig.de/∼autotyp

multiple branches, the migrations into South Asia, and the strong cultural ties to and within the Ancient Near East.

• Southeast Asia, including insular Southeast Asia up to the Wallace Line, delimited in the west by the western slopes of the Pātkai and Chin ranges and then, following the eastern limits of the Tibetan Plateau up to about 36° N and then across the plains to the Yellow Sea. Therefore, Southeast Asia includes the Yunnan and Sichuan hills, regions that played a vital role in the population history of Southeast Asia (cf. Enfield 2005 for recent discussion of Southeast Asian areality). This area is compared to both the rest of the world and to the rest of Eurasia, including Europe.

• The Eurasian enclaves: the Himalayas and the Caucasus. The enclaves are located at the fringes of the major Eurasian spreads through steppe, through Southwestern and South Asia, and through Southeast Asia. This location leads us to expect that the regions will deviate from the surrounding spread areas in their typological profile, an expectation confirmed by an earlier study (Bickel and Nichols 2003). Since the enclaves are part of Eurasia, our survey compares them to the rest of Eurasia (but not to the rest of the world).

• Sahul: Southern Australia and highland and southern New Guinea. This discontinuous area is demonstrated by Nichols (1997a) and Maddieson (forthcoming). It appears to be the result of coastal settlement activity bringing in distinctive types around the area, which had been settled earlier by the ancestral Sahul population. Our test compares Sahul to the rest of the world.

• Australia. Apart from exhibiting Sahul areality (see just above), it is possible that Australia as a whole is areally distinct from the rest of the world.

• The Circum-Pacific (CP) area: the Americas, Australia, New Guinea, Oceania, and the north Asian coast. This large and old area is motivated by the general population history of the Americas, as reflected for example by the Haplogroup B distribution (for some recent discussion, see Bickel and Nichols 2005b, 2006). The area is continuous in the sense that no other area splits it, though many of its parts are separated by water. Note that Southeast Asia is not part of the CP area because its current linguistic population shows strong historical continuities with more interior Asian regions throughout China and South Asia.

31.3 Variables surveyed

We define *case* as overt marking of the syntactic or semantic function of a nominal or pronominal argument or adjunct, marked on the nominal or pronominal itself or its dominating N′ or NP by an affix, clitic, or case word (by which we

978–0–19–920647–6 32-Spencer-ch31 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 482 of 493 May 16, 2008 18:59

mean a morpheme which is an independent word phonologically and prosodically but not syntactically, i.e. it is not an adposition, relational noun, or other word with syntactic properties such as licensing or assigning case, triggering agreement, heading a phrase, etc.). Some variables, however, gloss over the case vs. adposition distinction, surveying any kind of dependent marking of arguments in a clause or NP. We note this in the prose for those variables below. A zero-marked case category in an otherwise overtly marked case paradigm is counted as a case and as having the properties of its overt counterparts (examples are the nominative singular of Turkish or the genitive plural of Russian in some declensions).

In the following we give brief descriptions of all variables, beginning with those that concern the basic presence vs. absence of cases and then proceeding to their morphological form, behaviour, syntax, and alignment. Space limits prohibit more extensive discussion or exemplification, but all variables are published and we indicate the relevant source.

Some of the variables taken from WALS have been recoded in order to target specific issues, e.g. the base frequency of accusative vs. any other alignment (cf. explanations in section 31.4 below). Some of the separately listed variables have similar substance but are differently defined (e.g. one binary and one four-way, one from Autotyp data and one from WALS data, etc.), as we cannot know in advance which structural patterns are most strongly affected by biogeographical factors. Each variable is numbered here according to the order of its presentation below. In the prose (but not in tables), these variable numbers (and no others) are always in square brackets. In the listing below each variable is given a name that is mnemonic but not a full grammatical characterization of the variable.

31.3.1 Presence of cases

[1] Presence of cases/adpositions. Binary: presence vs. absence of cases/adpositions. Source: Dryer (2005b), recoded as value 9 ('no cases/adpositions') vs. all others.

[2] Number of cases (nouns only). Scalar: values run from zero to 10+ (the number of distinct cases in any one number paradigm). Source: Iggesen 2005c, recoded with 'zero' and 'exclusively borderline case marking' (value 9) grouped as Ø.

31.3.2 Morphological form of cases

Fusion, position, exponence, and flexivity are discussed in Bickel and Nichols (2007), and two of these (fusion and exponence) are also discussed in Bickel and Nichols (2005a, 2005c, respectively).

978–0–19–920647–6 32-Spencer-ch31 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 483 of 493 May 16, 2008 18:59

31.3.2.1 *Fusion*

Fusion is defined in strictly phonological terms and assesses whether, and if so how, case markers show phonological interaction with their host. We distinguish between the absence of such interaction ('isolating'), concatenative affixal marking, and marking on or in stems, including tonal ablaut (also cf. Bickel and Nichols 2005c). Phonological interaction can go together with morphological integration ('affixes'), but it can also affect independent syntactic units (e.g. when case adpositions phonologically cliticize to their host).

[3] Fusion. Binary: Phonologically isolating vs. other case marking.

[4] Fusion. 3-valued: Phonologically isolating vs. concatenative effects vs. tone. (The variable also distinguishes additional values, such as (pure) stem ablaut, replacive markers, prosodic word form templates, or reduplicative morphology, but our sample of case markers only contains isolating, concatenative, and tonal marking.)

31.3.2.2 *Position*

Whether the case marking precedes, is internal to, or follows its host. Position covers both independent grammatical words (adpositions) and case affixes.

[5] Position: preposed. Binary: Preposed vs. other.

[6] Position. 3-valued: Preposed vs. postposed vs. internal (ablaut apart from tonal ablaut, which was not included).

31.3.2.3 *Exponence*

Expression of categories other than case in the same morpheme as case. In Bickel and Nichols (2005a) and subsequent survey work we found case to share exponence only with number, tense/aspect/mood (TAM), or referential specification about topicality, definiteness, or specificity. Case + TAM coexponence is attested in only two languages, Logbara (Central Sudanic) and Kayardild (Tangkic) (see Crazzolara 1960 and Evans 1995a respectively for the primary analysis and Bickel and Nichols 2005a for a summary presentation). The pattern is not included in any further tests here. What we count as case + reference coexponence are cases that are obligatorily used, i.e. where speakers do not have a choice of not using a case marker. An example is the nominative in Tagalog which must be assigned to one of the arguments and entails topicality of that argument. The distribution of differential case-marking, where speakers have a choice of using vs. not using a particular case marker, is likely to be very different from that of case + reference coexponence, but we are not aware of a world-wide and sufficiently sampled database on differential case-marking. Indo-European languages are sometimes analysed as

978–0–19–920647–6 32-Spencer-ch31 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 484 of 493 May 16, 2008 18:59

showing case + gender coexponence, but in all cases we surveyed, this is actually the result of declension class allomorphy in case markers.

[7] Mono/polyexponence. Binary: monoexponential marker (case only) vs. polyexponential (case plus one or more other categories).

[8] Case + number. Binary: case + number coexponence vs. all other case forms.

[9] Case + reference. Binary: case + reference coexponence (case plus definiteness, case plus specificity, etc.) vs. all other.

31.3.2.4 *Flexivity*

Allomorphic alternation in the morpheme(s) marking one and the same case category, governed by either lexical stem classes (i.e. declension classes) or cross-classifying categories (e.g. different case allomorphs on nouns vs. pronouns or different plural allomorphs depending on whether there is also an article or possessive agreement, etc.)

[10] Flexivity. Binary: some flexivity in the case morphemes vs. none.

31.3.2.5 *Tone*

Case marked by tone.

[11] Tonal case. Binary: one or more of the cases is marked exclusively by tonal ablaut vs. all other. Source: merged Autotyp and WALS (Dryer 2005b) datasets (which exhibit complete agreement in coding).

31.3.3 Morphological behaviour of cases

31.3.3.1 *Host*

These variables have to do with whether the case morpheme is restricted to a specific stem class, i.e. a true affix, or whether it is unrestricted and attaches to phrases of variable category, i.e. a true clitic. Because phrasal category structure and the exact distribution of case markers are often underanalysed in grammars (especially older grammars), the coding differs in Autotyp and WALS to a degree that merging the datasets is unwarranted.

[12] Restricted case A. Binary: case is selectionally restricted to the stem of the head word in the argument phrase vs. positioned relative to the phrase boundary and compatible with hosts of any category.

[13] Restricted case W. Same variable; source: Dryer (2005b), recoded as binary with values 1–5 (various kinds of affixation) vs. 6–8 (phrasal clitics); value 9 (absence of case) removed.

978–0–19–920647–6 32-Spencer-ch31 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 485 of 493 May 16, 2008 18:59

31.3.3.2 *Spreading*

These variables have to do with whether case marking spreads from the head noun to other words in the NP, e.g. to an agreeing attributive adjective or otherwise ends up not on the head but on another word or at the NP boundary.

[14] Case spreading. 4-valued: Case spreads vs. is limited to phrase-initial vs. to phrase-final vs. to the head.

[15] Case spreading. Binary: Case does vs. does not spread.

31.3.4 Core cases

[16] Core cases. Binary: no overt case marking on core arguments (S, A, O) vs. some case marking on one or more core arguments.

[17] A ≠ O. Binary: case or other dependent-marking differentiates A from O vs. no differentiation (either no marking or neutral alignment). Source: merged from information in Autotyp and WALS (Comrie 2005).

[18] A case. Binary: the A does vs. does not have case or other dependent marking.

[19] O case. Binary: the O has vs. does not have case or other dependent marking.

[20] A and/or O case. Binary: A and/or O does vs. does not have case or other dependent marking.

31.3.5 Alignment

31.3.5.1 *Noun alignment, i.e. overt alignment of noun case inflection*

Variables [21–25] have to do with the alignment of S with A or O; [26–27] have to do with alignment of objects (direct/indirect vs. primary/secondary, following Dryer 1986).

[21] Accusative/other. Binary: accusative vs. other alignment (excluding neutral alignment, i.e. S=A=O and caseless languages). Source: merger of data from Autotyp and Comrie 2005.

[22] Alignment A + W. 4-valued: accusative vs. ergative vs. three-way (S≠A≠O) vs. split (part S=A and part S=O). Source: merged data from Autotyp and Comrie 2005.

[23] Alignment A. Same; source: Autotyp only.

[24] Alignment W. Same; source: Comrie (2005) only.

[25] Alignment 5. 5-valued: Accusative vs. ergative vs. marked accusative vs. three-way vs. split. Data source: Comrie (2005).

[26] PO alignment. Binary: primary object alignment vs. other (excluding neutral alignment and caseless languages).

[27] Object alignment. 4-valued: direct object vs. primary object vs. three-way vs. split.

978–0–19–920647–6 32-Spencer-ch31 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 486 of 493 May 16, 2008 18:59

31.3.5.2 *Pronoun alignment, i.e. overt alignment of pronoun case inflection*

These variables are otherwise identical in definition and ordering here to those for nouns in §31.3.5.1. [28–32] have to do with subject alignment, [33–34] with object alignment.

[28] Pro accusative/other. Binary: accusative vs. other alignment (excluding neutral alignment, i.e. S=A=O). Source: merger of data from Autotyp and Comrie (2005).

[29] Pro alignment A + W. 4-valued: accusative vs. ergative vs. three-way (S≠A≠O) vs. split (part S=A and part S=O). Source: merged data from Autotyp and Comrie (2005).

[30] Pro alignment A. Same; source: Autotyp only.

[31] Pro alignment W. Same; source: Comrie (2005) only.

[32] Pro alignment 5. 5-valued: Accusative vs. ergative vs. marked accusative vs. three-way vs. split.

[33] Pro PO alignment. Binary: primary object alignment vs. other (excluding neutral alignment and caseless languages).

[34] Pro object alignment. 4-valued: direct object vs. primary object vs. three-way vs. split.

31.3.6 Syncretism in cases

Syncretism in cases involves the identical coding of argument roles in some but not other lexically defined paradigm classes. If coding is identical in core roles only (i.e. S and either O or A) this amounts to a lexical split in alignment, where one lexically defined paradigm class codes S like O, while the other codes S like A. However, because of the small number of relevant examples, we limit the survey to the base frequency of any kind of syncretism.

[35] Syncretism. Binary: syncretism (between different paradigms) in one or more cases vs. none. Source: Baerman and Brown (2005), recoded to conflate values 2 (syncretism in only core cases) and 3 (syncretism in core and non-core cases) and removing 1 (no case).

31.4 Methods

A number of the variables described above are coded in both Autotyp and WALS. Where there is parallel coding, we checked the coding for consistency. In those cases

where coding was consistent to 95 per cent or higher, we merged the datasets, giving preference to the Autotyp coding in the handful of mismatches (variables [5], [6], [11], [17], [21], and [28]). Where coding consistency was higher than 90 per cent, we report results of the individual as well as the merged datasets (variables [22] and [29]). Where coding was less than 90 per cent consistent we did not merge datasets. This was only the case with the variables coding host restrictions, i.e. those capturing the notoriously difficult clitic vs. affix distinction (variables [12] and [13]).

The Autotyp datasets are all pre-sampled for genealogical balance, i.e. they include at most one representative per major branch.[2] The WALS and the merged datasets are not pre-sampled. In order to control for inflationary effects of possibly over-represented stocks, we performed a genealogical distribution analysis following Bickel 2007d. In this analysis, each variable is first cross-tabulated with each areal predictor variable (e.g. presence of case * Eurasia, etc.). Then, for each cell in these tables, we tested whether any of the genealogical levels registered in our database (stock, major branch, subbranch, sub-subbranch, lowest subbranch; cf. Nichols forthcoming, b) shows statistically significant skewing, e.g. nine out of ten languages from the same group showing case. If this is so, the number of languages with the same typological value in the same genealogical group is reduced to one, assuming that the skewing could be caused by shared inheritance instead of the areal factor of interest. If there is no skewing in a given genealogical group, all languages from that group are included in the sample. For the one scalar variable (number of cases: [2]) we performed pre-sampling, randomly selecting one representative of each major branch, because we have currently no good algorithm at our disposal that would allow sampling of interval distributions based on genealogical distribution analysis.

The genealogically pre-sampled or genealogically controlled tables were then submitted to significance testing, applying the distribution-free methods proposed in Janssen et al. (2006). For categorical variables, we used the Fisher Exact Test, and where the dataset was too large for an exact computation of p-values, we relied on the approximation methods for this test implemented in the statistical software package R (R Development Core Team 2006). The one scalar variable in our survey ([2], number of cases) was tested in a randomization-based analysis of variance, as developed by Janssen et al. 2006.

Since we simultaneously tested each typological variable for ten possible area effects, we needed to control for familywise error, i.e. for the risk that some variables reach significance because of the sheer number of possibly significant associations. In order to control for this, we applied Holm corrections to each set of ten test results (Holm 1979, as implemented in R). This yields fairly conservative p-values, and we therefore set the level of rejection at .05.

[2] A complete listing of our genealogical sample is available on the Autotyp web site (www.uni-leipzig.de/~autotyp).

Table 31.1 reports the p-value of each test but also indicates those tests that reached significance before the Holm adjustment was applied (marked by asterisks). These call for further analysis of richer samples and follow-up research on the relevant areal hypotheses.

31.5 Results and discussion

Table 31.1 shows the typological variables, the areas, and all frequency differences that were statistically significant at $p < .05$. All of the statistical comparisons pit the macroarea, area, or enclave in question against the rest of the world's languages, unless otherwise indicated. There is no appreciable difference between binary and larger-valued variables in their significance levels or the number of times they yield significance, and this suggests that our results do not depend on the arity of variables (indicated in the column 'Values' in Table 31.1).

About half of the variables (18/35) yield significance somewhere, most of them in more than one area. Those that do chiefly have to do with presence and number of cases (variables [1], [2], [16], [18–20]); the clearest aspects of morphological form, namely fusion (isolating vs. concatenative vs. tonal: variables [3–4]) and position (preposed vs. postposed, etc: variables [5–6], [11]); and alignment, chiefly subject alignment (variables [22–25], also [28–32]). (The absence of significant areal effects in object alignment could well be due to the low number of datapoints in our sample, and therefore our results do not support any strong inference in this case.)

Case presence (variable [1]) and the number of cases ([2]) show frequency peaks in Eurasia (59% of the languages have case, mean number: 4.6) and Australia (81%, mean: 6); the number of cases is significantly depressed in Southeast Asia (28%, mean: 2.4) and, with a borderline effect, Africa (45%, mean: 2.5).[3] (The significant effect of CP*[1] in Table 31.1 is the indirect result of the frequency depression in Southeast Asia and Africa, which lowers the total frequency in the non-CP languages.) Case or other dependent marking (variables [18–20]) for A and/or O again show frequency peaks in Eurasia (86% for variable [20]) and Australia (80%). The Circum-Pacific region shows significantly less dependent marking (49%) on A and/or O, preferring instead head marking (.Nichols 1992). (Maps on the presence of cases are available in WALS.)

[3] For the datasets that were genealogically sampled based on a distribution analysis within genealogical taxa (cf. section 31.4), we report percentages and odds from the genealogically balanced numbers (not the raw data) since (a) these were also the numbers that entered the tests, and (b) these are the only numbers that are comparable to the genealogically pre-sampled datasets in Autotyp.

Table 31.1. Results of testing structural variables for areal distributions

No.	Variable	Values	Source	N	Africa	Europe	Eur:Asia	Eurasia	SEA	SEA:Eurasia	Enclaves	Australia	Sahul	CP
1	Cases	2	W	934	*	*		0.0179	0.0171		*	*	*	0.0001
2	Number of cases	scalar	W	199	*			0.0224		0.0020		0.0036		
3	Fusion	2	A	133					0.0046	0.0034				
4	Fusion	3	A	133	0.0024				0.0097	0.0030				*
5	Position: Preposed	2	A&W	636				0.0063	0.0115		*			0.0205
6	Position	3	A&W	622				0.0225	0.0239		*	0.0164	0.0225	0.0280
7	Mono/polyexponence	2	A	137										
8	Case + number	2	A	137										
9	Case + reference	2	A	137										
10	Flexivity	2	A	72								*		
11	Tonal case	2	A&W	634	0.0005									0.0053
12	Restricted case A	2	A	72										*
13	Restricted case W	2	W	596										
14	Case spreading	4	A	63										
15	Case spreading	2	A	63										
16	Core cases	2	A	239	*			0.0002						
17	A ≠ O	2	A&W	383										
18	A case	2	A	259				0.0000		0.0252		0.0077		0.0252
19	O case	2	A	256				0.0000		*		*		0.0009

(*cont.*)

978–0–19–920647–6 32-Spencer-ch31 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 489 of 493 May 16, 2008 18:59

Table 31.1. (*Continued*)

No.	Variable	Values	Source	N	Africa	Europe	Eur:Asia	Eurasia	SEA	SEA:Eurasia	Enclaves	Australia	Sahul	CP
20	A and/or O case	2	A	256				0.0000		*		0.0491		0.0040
21	Accusative/other	2	A&W	190	*							*	0.0028	
22	Alignment A + W	4	A&W	82	0.0125							0.0076	0.0047	
23	Alignment A	4	A	114	0.0313						*	0.0216		*
24	Alignment W	4	W	92	*						*		*	0.0075
25	Alignment 5	5	W	92	0.0171			*			*			0.0144
26	PO alignment	2	A	48										
27	Object alignment	4	A	48								*		
28	Pro accusative/other	2	A&W	202	0.0141								*	
29	Pro alignment A + W	4	A&W	91	0.0177							0.0368	*	
30	Pro alignment A	4	A	125							*			*
31	Pro alignment W	4	W	90										
32	Pro alignment 5	5	W	90										
33	Pro PO alignment	2	A	42										
34	Pro object alignment	4	A	42										
35	Syncretism	2	W	74										

Figures report adjusted p-values of significant results; * indicates distributions that were significant before application of the Holm correction. Under 'Values' we list the number of possible values of each variable (its 'arity'); A = data from Autotyp, W = from WALS, A&W = merged data from both sources. N gives the number of languages that entered the analysis (including both the genealogical analysis and the hypothesis testing, as described in section 31.4). The data behind Table 31.1 is available online at the Autotyp web site: www.uni-leipzig.de.

978–0–19–920647–6 32-Spencer-ch31 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 490 of 493 May 16, 2008 18:59

Table 31.2. Prefixal case marking

	preposed case	case not preposed	Sum
Southeast Asia	6 (60%)	4 (40%)	10
Rest of the world	29 (14%)	185 (86%)	214
Sum	35	189	224

Source: genealogically balanced sample from merger of Autotyp and WALS (Dryer 2005b).

The fusion variables ([3] and [4]) show a strong and well-established preference for isolating case morphology in Southeast Asia where isolating case morphology rises to a 5-to-5 tie (contrasting with a 16.4-to-1 chance in the rest of the world and 31-to-0 in the rest of Eurasia). Tonal case marking (variable [11]) is represented only in Africa (with 6 examples: in Maba, the Nilotic languages Maa, Nandi, and Shilluk, and the Benue-Congo languages Nande and Yoruba).[4] Perhaps it is a singularity unique to Africa (like phonemic clicks), or perhaps it occurs there simply because Africa is the only macroarea where both tones and cases occur with any frequency in the same languages. (Exclusively tonal marking of other inflectional categories, namely those of verbs, does occur in the Iau language of New Guinea: Bickel and Nichols 2005c: 88, 89. It possibly occurs elsewhere, as the sample for that survey was not large.)

The position variables ([5] and [6]) reveal a strong (9.8-to-1) universal dispreference against preposed case, but this trend is reversed in Southeast Asia, where prefixed or procliticized cases outnumber other cases by 6 to 4; cf. Table 31.2. All other significant effects of variables [5] and [6] in Table 31.1 are a side effect of this since, on the one hand, Eurasia contains Southeast Asia, artificially increasing the number of preposed cases in Eurasia, and, on the other hand, all other areas were tested against the rest of the world, which again included Southeast Asia. Case marking by stem-internal ablaut alone does not reach appreciable frequencies anywhere. There are seven such cases attested in total in the combined Autotyp and WALS dataset, four from Australia (Anindilyakwa, Bunuba, Nyulnyul, Yawuru) and three from Africa (Dinka, Nuer, Yoruba).

Alignment patterns (variables [21–34]) show increased frequencies of ergativity in Australia (92%) and Sahul (75%, plus 10% split). There are also appreciable frequency increases of ergativity in the Circum-Pacific (46% in the Autotyp sample) and the Eurasian enclaves (56%), but the statistical effects of this are borderline: the effects are only clearly present before controlling for multiple testing and they

[4] The significant effect of CP*[11] is again an artefact because the non-CP languages contain the tonal cases of Africa.

Table 31.3. Marked nominatives

	marked nominative	other	Sum
Africa	4 (36%)	7 (64%)	11
Other	2 (3%)	63 (97%)	65
Sum	6	70	76

Source: genealogically balanced sample from WALS (Comrie 2005).

disappear when the sample becomes richer (i.e when the Autotyp and WALS samples are merged). In Chapter 20 (this volume) we propose that alignment patterns need to be surveyed relative to lexical classes and valence structure, and in general require much more empirical groundwork than the summary data provided in current databases. Future research will show whether the Circum-Pacific and enclaves reveal robust distributional differences once alignment patterns are surveyed in more detail.

Variable [25] (from Comrie 2005) distinguishes accusative alignment with an unmarked nominative vs. accusative alignment with a marked nominative. Marked nominatives are rare worldwide, but the presence of four such cases in Africa (Middle Atlas Berber, Igbo, Murle, and Oromo) is statistically significant despite the small numbers of cases in Africa in general. This is confirmed by a post-hoc analysis, testing marked nominatives against all other alignments in Table 31.3.

The significance effect of [25] *Circum-Pacific in Table 31.1, by contrast, results from different distributions of ergative and split marking; it is not confirmed by a post-hoc analysis (Fisher Exact test, $p = .39$). (The two non-African languages with marked nominatives in the sample are Maricopa in North America and Aymara in south America.)

Case spreading (variables [14–15]), flexivity (variable [10]), and syncretism (variable [35]) show no area signals, perhaps because the datasets for these variables are so small. However, these features are also known to be fairly resistant to areal spread, and it is therefore likely that the absence of significant test results points to the absence of macro-areal distributional skewings. Another variable which is likely to resist spread, but which is better surveyed, is case exponence (variables [8] and [9]). Any kind of polyexponence is relatively rare worldwide, showing up in only 13 per cent of the languages surveyed ($N = 137$). In a pattern too localized to show up at the level of macroareas in Table 31.1, case + number coexponence occurs in all morphologically conservative languages of the Uralic and Indo-European families (Bickel and Nichols 2005a).

31.6 Conclusions

Our overall conclusion is that the worldwide geographical distribution of case inflection and its various grammatical properties mirrors global linguistic geography reasonably well and indicates that aspects of position and fusion of case markers, their presence vs. absence, and their alignment are prone to areal spread while aspects of exponence, flexivity, syncretism, and phrasal behaviour tend to resist spread.

Most geographical areas tested show some distributional effect, which is however boderline for the Eurasian enclaves. Only one area did not reveal any distributional effect: Europe. In Table 31.1, there is borderline indication for increased frequencies of cases (variable [1]) but this recapitulates the increased frequency in the larger Eurasian area, as shown by the fact that testing Europe against the rest of Eurasia did not reveal any significant difference. Thus, in none of the variables surveyed is Europe significantly different from either the rest of the world or just Asia. Statistically speaking, the overall strong similarity of Europe and Asia makes their composite macroarea Eurasia both large enough and monolithic enough to reach significance on several variables. Linguistically speaking, the similarity of Europe to Asia supports – as far as case marking is concerned – the decision to regard Eurasia as a single macroarea. Historically speaking, it confirms that Europe is the endpoint of the population movements along the Eurasian steppe, which is not separated from it by any major discontinuity.

Acknowledgements

The Autotyp database project has been supported in part by Swiss NSF grants 08210-053455 (1998–2001, Bickel) and 610-0627 (2001–2002, Bickel), German DFG grants BI 799/2-1 (2003–2005, Bickel & Hall), BI 799/2-3 (2005–2007, Bickel), BI 799/3-1 (Bickel & Bornkessel), and U.S. NSF grants 92-22294 and 96-16448 (Nichols). We thank two anonymous reviewers for helpful comments on an earlier draft.

CHAPTER 32

CASE AND CONTACT LINGUISTICS

LARS JOHANSON

32.1 COPYING CASE MARKERS AND CASE FUNCTIONS

LANGUAGE contact affects case categories in various ways. This chapter is concerned with effects of contacts between linguistic codes (languages, unrelated or related, or language varieties): changes in one code on the model of another. It deals with inflectional case markers, affixes, and adpositions from which they evolve. Though most adpositions express more specific relations, some are relatively desemanticized. Affixes and case-like adpositions may fulfil similar functions; cf. the close correspondences between Dravidian case suffixes and Indic postpositions.

Case markers and case functions are acquired through what is called 'borrowing', 'diffusion', 'transfer', 'interference', 'replication', etc. Speakers copy case markers or case functions from a model code (a 'source', 'donor', or 'diffusing' language) and insert the copies into their basic code (a 'recipient' or 'replica' language). The term 'copying' is preferred here to stress the non-identity of models and copies.

978–0–19–920647–6 33-Spencer-ch32 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 495 of 502 May 16, 2008 18:59

32.2 Global copying: 'transfer'

Case markers can be copied *globally*, as a whole, including their material shape and properties of meaning, combinability, and frequency, a process often called 'transfer'. The copy is inserted into a position felt to be equivalent to that filled by the original in the model code. According to Thomason and Kaufman, copying of case affixes and case categories is possible 'under strong cultural pressure'. New case affixes may be added to native words, 'especially if there is a good typological fit in both category and ordering' (1988: 75).

Case affixes are relatively seldom copied globally. The Lithuanian illative suffix *-n* is considered to be a copy of an illative suffix of an older Finnic language. The Kurmanji (NW Iranian) locative marker *-de*, e.g. in *memleketde* 'in the homeland', might be a copy of Turkish *-de*, or a remnant of a native postposition (Bulut 2006). As a rule, only single markers are copied: 'the transfer of a full grammatical paradigm ... has apparently never been recorded' (Weinreich 1953: 43–4). However, Northern Tajik (SW Iranian) varieties, spoken in Central Asia, have copied the Uzbek core case suffixes *-gä* dative, *-dä* locative, and *-dän* (ablative). Resígaro (Arawak), spoken in Peru, has copied several oblique cases from Bora (Witotoan) (Aikhenvald 2001).

The case systems of fusing languages such as Latin provide few suitable models for copying. In older learned texts in European languages, Latin nouns were often copied globally together with their case markers, e.g. dative *Christo*. After non-Latin prepositions, they assumed the case markers required by the corresponding Latin prepositions, e.g. German *mit den pronomibus* 'with the pronouns'.

Case-like adpositions are more easily copiable. Irano-Turkic varieties have copied simple Persian prepositions, e.g. Khalaj *bī sän* 'without you'. The Kurmanji preposition *gor(a)* 'according to' is copied from the Turkish postposition *göre*; Mari *köra* 'because of' from Tatar (NW Turkic) *körä*. Karaim (NW Turkic) has copied *okolo* 'around, about' from Slavic, and *puk'i* from Polish *póki* 'up to, until'. A global copy of a periphrastic Latin preposition is German *in puncto* 'in respect of'. Even whole case forms can be globally copied, e.g. English *theim* 'them' from Scandinavian replacing Middle English *him* (Baugh and Cable 1993: 102, Morse-Gagne 2003).

32.3 Selective copying: 'replication'

Copying does not necessarily involve morphological material. Properties of case markers – semantic, combinational, and frequential properties – may be copied

978–0–19–920647–6 33-Spencer-ch32 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 496 of 502 May 16, 2008 18:59

selectively from one code into another without material shapes, the means for expression being provided by the basic code. This process is called 'replication', 'semantic interference', 'loan translation', 'calquing', 'indirect diffusion', etc.

On the basis of structural and conceptual similarities, an equivalence relation is established, consciously or intuitively, between a marker in a model code and a suitable target in the basic code, a native segment onto which the relevant properties can be copied. Targets – affixes, adpositions, etc. – are reanalysed with respect to these properties, their functions becoming more similar to those of the models.

A basic code may have a comitative marker, whereas its model code equivalent fullfils both comitative and instrumental functions. Copying properties from the latter marker allows the use of the former marker in both functions. Speakers of Basque have copied properties of Romance models onto the comitative case suffix *-ekin* to give it an additional instrumental function (Stolz 1996b). Karaim *-Bɔ* 'with' displays uses resulting from copying of semantic and combinational properties of the Russian instrumental case (Csató 2000). Similar processes are observed for case-like adpositions. English semantic-combinational influence on Guernésiais, the Norman variety of Guernsey, is seen in the trend to use one single form corresponding to *with* instead of separate comitative and instrumental forms. The instrumental preposition *atou* is vanishing in favour of the comitative *dauve* (Jones 2002: 157).

32.4 Results of copying

Copying of case markers and functions leads to various results. Semantic copying affects the meaning, combinational copying changes the applicability to contexts, and frequential copying leads to increased or decreased occurrence. There are always corresponding differences between models and copies. Copies are often subject to creative restructuring according to general cognitive and communicative principles.

The ultimate outcome is functional reorganization of the basic code: modification of features, addition of features, loss of features. An existing case marker assumes modified functions, a new case marker is added, or an existing case marker is eliminated. There may be shorter or longer periods of competition between conflicting markers.

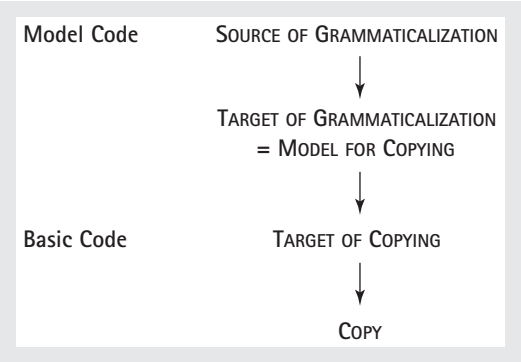


Figure 32.1.

32.5 Selective copying and grammaticalization

The relationship between selective copying and grammaticalization is based on the following distinctions: (1) the source of grammaticalization of the model code marker, (2) the target of grammaticalization of the model code marker, which is identical to the model for copying, and (3) the target of copying (see Figure 32.1).

The interrelations may be illustrated with a process found in Tariana (North Arawak) and described by Aikhenvald (2001). Though Arawak languages typically do not employ case marking for core grammatical relations, Tariana has developed, on the model of the East-Tucanoan suffix *-re*, a core case marker for topical and specific referents in non-subject functions. This is the result of reanalysis and reinterpretation of the target *-naku/-nuku*, an allative marker meaning 'on/to the surface'. Since grammatical means for this case-marking are supplied by a previously existing Arawak structure, no grammaticalization is involved. The source of grammaticalization of the East-Tucanoan marker is irrelevant for the copying process. What is copied is not the history of the model code marker, but the result of a code-internal process at a specific stage of grammaticalization.

Copying is not a grammaticalization process. Grammaticalization proceeds unidirectionally from less to more grammaticalized items. Fresh copies, however, mostly represent less advanced stages than their models. Their use is often pragmatically determined, contextually restricted, and optional rather than obligatory. Heine and Kuteva, who take changes in case functions to follow the same principles of grammaticalization as changes not involving language contact, note: 'wherever there is sufficient evidence, it turns out that the replica construction is less grammaticalized than the corresponding model construction' (2005: 101). If we were to

take the copying act itself to be a grammaticalization process, such phenomena would have to be viewed as instances of reverse directionality, i.e. violations of the unidirectionality principle. I do not share Heine and Kuteva's opinion that grammatical replication involves a grammaticalization process.

Copies are immediately subject to internal processes of the basic code, where they may reach more advanced stages of grammaticalization. Again, this is not a result of the copying itself, but a matter of code-internal development. Copies at early stages of grammaticalization may go through various new stages, whereas copies at advanced stages have few chances to develop further (except eroding to zero).

32.6 Copiabililty and stages of grammaticalization

Grammaticalization of case markers normally follows the path LEXICAL ITEM > CASE-LIKE ADPOSITION > CASE AFFIX > ZERO. The transition from adpositions to case affixes is characterized by extension of occurrence, desemanticization, decategorialization, and erosion. Adpositions denoting location, source, and destination often develop into case affixes.

Copiability ('borrowability') is correlated with the stage of grammaticalization as reflected in degrees of saliency of meaning and shape. Old case markers with central syntactic functions are less transparent and less salient, mostly displaying reduced shapes. Their opaque structure makes it difficult to decide whether they are copies or not. Case-like adpositions, which are more transparent and open to analysis, with materially richer shapes and relatively specific meanings, are more copiable. Transparency makes it easier to find a natural corresponding target of copying. The Tajik postposition *dīdä* 'because of' is a semantic-combinational copy of Uzbek *körä*. Properties of *kör-* 'to see' plus a converb suffix have been copied onto an item formed from the past stem of Tajik *dīdän* 'to see'. Under Yakut (Turkic) influence, dialects of Even (Tungusic) tend to replace the directive-locative marker *-kla* by a construction with the postposition *istala*, a converb of *is-* 'to reach' (Malchukov 2006b). Elaborate, periphrastic adpositions representing early stages of grammaticalization and employing nominal items as the semantic core, e.g. *à cause de*, *by reason of*, are easy to analyse and copy into languages that possess equivalent nouns. Irano-Turkic varieties have copied periphrastic Persian prepositions, e.g. *äz täräf-e* 'on the part of'. Such complexes are grammaticalized further, e.g. English *because of*, which still reflects a structure containing the copied French noun *cause*. The Uralic languages of Europe have developed their case systems substantially through

978–0–19–920647–6 33-Spencer-ch32 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 499 of 502 May 16, 2008 18:59

grammaticalization of postpositions. Several Hungarian cases, the Estonian comitative, etc., have been formed relatively recently.

Selective copying is facilitated by congruent structures. Copied case affixes fill positions corresponding to those of their models. Copied adpositions may occur in a different linear order relative to their head, e.g. the Basque postposition *kontra* as a copy of the Spanish preposition *contra* 'against'. Karaim and Irano-Turkic varieties may use copies of prepositions as postpositions (Csató 2000).

32.7 Copying valency patterns

Copied combinational properties may affect rules for case assignment. Model code predicates may trigger copying of their valency patterns for basic code equivalents. Turks in Western Europe tend to choose the accusative (instead of dative or ablative) with *sor-* 'to ask', influenced by equivalent foreign verbs taking direct object markers. Finnish schoolchildren in Sweden tend to use *käydä koulussa* (inessive) instead of *käydä koulua* (partitive); cf. Swedish *gå i skolan* 'to go to school' with the preposition *i* 'in'. Influenced by Russian verbs governing the instrumental, Uzbek speakers tend to overextend the use of the postposition *bilän* 'with', e.g. X *bilän qiziq-* 'to be interested in X' (*interesovat'sja* + instrumental) instead of dative (*-gä*) *qiziq-*. Copying of valency patterns may be particularly tempting when the model and the copy are phonetically similar. Due to identification of *to* and *for* with Swedish *till* and *för*, respectively, speakers of American Swedish have been observed to use *van till* instead of *van vid* 'accustomed to' and *för ett år* instead of *i ett år* 'for a year'.

32.8 Examples of selective copying

The sole function of Persian *-rā* (< *rādi* 'because of') is to mark specific direct objects, though it previously also had other functions. This similarity to the use of Turkic accusative markers might well be contact-induced (Johanson 2002: 102–3).

Turkish influence on the Ardeshenian dialect of Laz (South Caucasian), spoken in Northeastern Anatolia, has restructured the system of core case marking. Young urban Turkish–Laz bilinguals extend the use of the goal marker *-ša* to mark indirect objects, according to the use of the Turkish dative suffix *-(y)A* (Haig 2001: 214–16).

978–0–19–920647–6 33-Spencer-ch32 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 500 of 502 May 16, 2008 18:59

Ergative systems are often susceptible to contact-driven change. Under Azeri and Armenian impact, Udi (Northeastern Caucasian) has given up its ergative construction. The Mingrelian ergative marker was once extended to cover both the subject of intransitive clauses and the subject of transitive clauses in the aorist, resulting in a nominative–accusative system (Dixon 1994: 202). Bulut (2006) notes that the split ergative system of Diyarbakır Kurmanji has broken down in certain narrative structures, with nominative–accusative widely used instead of ergative–absolutive, probably due to internal tendencies supported by frequential copying from Turkish.

Influenced by Georgian (Southern Caucasian), Ossetic (Eastern Iranian) has added new agglutinatively arranged case markers to the inherited Indo-European system. Instead of the Old Armenian fusional case-marking system, modern Armenian uses, for the same categories, native forms with selectively copied functions arranged agglutinatively on the model of Turkish.

Uralic substratum influence ('imperfect learning') is assumed to have influenced Slavic and Baltic, e.g. giving rise to partitives. The new Lithuanian illative, allative, and adessive cases were developed from postpositions, probably under Balto-Finnic influence. The use of the nominative for objects in Russian impersonal constructions may have been copied from Finnish (Timberlake 1974). The predicative instrumental construction used in most Slavic languages and Lithuanian, e.g. Russian *On byl soldatom* 'He was a soldier', may be due to early Uralic influence. The corresponding Finnish case is the essive (Veenker 1967: 131). Karaim (Turkic) has copied the pattern from Slavic and Lithuanian, e.g. *Ol ed'i yavanbɔ* 'He was a soldier', with the instrumental marker *-Bɔ*.

Whereas Indo-European languages have generally suffered substantial loss of cases, most Slavic languages and Lithuanian have retained many inherited cases, possibly due to reinforcing and conserving frequential copying from the case-rich Uralic languages. The Finnish partitive has influenced the Russian use of genitive-marked direct objects with negations. Russian *-u*, once the genitive marker of a minor noun class, developed into a partitive marker when Uralic speakers shifting to Russian copied properties of their partitive onto it (Thomason and Kaufman 1988; Thomason 2001).

32.9 Polysemy, syncretism

Contact-induced extension of the use of case markers may cause polysemy; see the above-mentioned additional use of comitative markers in instrumental functions.

978–0–19–920647–6 33-Spencer-ch32 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 501 of 502 May 16, 2008 18:59

Merge of cases – one category extending its domain and taking over the function of another category – causes syncretism. Locatives are often used instead of datives. Yakut (Turkic) and the co-areal language Evenki (Tungusic) exhibit a dative–locative merger due to Mongolic influence. Merger of locatives and directionals is also a Balkan feature. Under the influence of corresponding markers in Macedonian, Albanian, and Serbian, the West Rumelian Turkish locative is generalized to express 'location' and 'motion toward', e.g. *bizde dʒel* 'come to us'. On the model of its Guernésiais equivalent *a*, the preposition *to* in Guernsey English is also used to express static location (Jones 2002: 147).

The use of datives or allative-goal markers is often extended to mark direct objects. A dialect of Kannada (Dravidian) has, under Indo-Iranian influence, extended the use of its dative to express human direct objects (Gumperz and Wilson 1971: 158). When communicating in Portuguese, speakers of Tariana use the allative-goal preposition *pra* 'to, towards, for' to mark definite and referential objects (Aikhenvald 2002: 314). A similar example is the direct object marking in Romance, e.g. Spanish, by means of the preposition *a* 'to'. The use of an allative–dative preposition as a direct object marker in Maltese (*lil*) and Spanish Arabic (*li*) may be influenced by this pattern (Heine and Kuteva 2005: 150–2).

32.10 Impoverished case systems

Case reduction is well known in immigrant contact situations, particularly in contacts with the highly reduced English case system. Clyne (2003: 124–30) summarizes recent studies on case loss and restructuring in immigrant varieties spoken in the USA and Australia. German varieties exhibit case syncretism, reduction to one common case or to two cases, nominative and oblique. Examples include loss of dative, instrumental, and locative in Polish, replacement of the Croatian dative and locative by the nominative and the accusative, respectively, omission of accusative, inessive, and superessive markers in Hungarian, case reduction in Finnish, e.g. loss of the partitive.

Extreme loss of case markers is found in makeshift languages, pidgins such as 'Gastarbeiterdeutsch' ('guest-worker German') or foreigner talk. In Russenorsk, once used by Russian merchants and Norwegian fishermen – the result of two-way copying processes between a simplified Norwegian variety and Russian foreigner talk – case relations were expressed with the all-purpose preposition *po*.

Impoverished case systems can be enriched by the creation of new case-like adpositions (see Kulikov, Chapter 28).

PART VI

INDIVIDUAL CASES: CROSS-LINGUISTIC OVERVIEWS

CHAPTER 33

TERMINOLOGY OF CASE

MARTIN HASPELMATH

33.1 Introduction

As in all areas of grammar, the terminology surrounding case phenomena is often not straightforward: Linguists with different backgrounds use the same terms for somewhat or radically different concepts, or they use different terms for very similar or identical concepts. It is unlikely that terminological consensus will emerge soon, primarily because there is no consensus about the concepts that we need, and terminological polysemy will continue to be rampant because there are many more concepts than handy terms. But it is useful to be aware of some of the most important terminological issues.

33.2 Basic notions

The term **case** can refer to an inflectional category-system (e.g. 'Many Australian languages have case') or to the individual inflectional categories or values of that system (e.g. 'Nhanda has seven cases'). In this respect, *case* behaves like other inflectional category-systems such as tense, aspect, mood, person, number, gender, i.e. we

are dealing with a systematic ambiguity that does not lead to misunderstandings. Along with the status of an inflectional category-system comes a range of old habits of talking, such as saying that a word-class '*inflects for* case', that a lexeme '*is in* the dative case', or that a form is the '*genitive plural*' of a lexeme. The latter expression not only illustrates a word order convention (we would not normally say 'plural genitive') and an abbreviatory convention (we would not say 'genitive case plural number'),[1] but also a third metonymic use of inflectional category words: they can also be used to denote **case forms**, i.e. words that express these categories (e.g. 'The genitive case of Latin *pater* is *patris*') (cf. Mel'čuk 1986: 37).

In an old terminology that is becoming obsolete, inflection for case is called **declension**, and a lexeme is said to **decline** when it changes its cases. This usage is the source of the term **indeclinable**, referring to words that do not show overt case distinctions although they would be expected to show them.[2] The term *declension* now mostly survives in the sense 'inflectional class defined by different case forms', a phenomenon that is best known from (especially the older) Indo-European languages.

The term *case* is from Latin *casus* 'fall(ing)', itself a loan translation from Greek *ptōsis* 'fall(ing)' (cf. loan translations in other languages such as German *Fall*, Russian *padež*, from *pad-* 'fall'). The idea seems to have been that of 'falling away from an assumed standard form' (Blake 1994: 19), and the terms *declension* (from *declinatio* 'turning away, deviation') and *inflection* (*inflectio* 'bending') are based on similar spatial metaphors for meaningful formal variations in the shapes of words.

Latin and Greek had five or six cases with relatively abstract syntactic–semantic functions, but linguists did not find it difficult to carry over the concept of case to languages with many more case distinctions (such as the Finno-Ugrian languages) or with rather different kinds of cases (such as the Australian languages). The function of cases is generally agreed to be that of 'marking dependent nouns for the type of relationship they bear to their heads' (Blake 1994: 1),[3] so that other nominal markings such as head marking for person, head marking for possessedness (*status constructus*), and NP marking for definiteness, topic, or focus have never been considered cases.

However, relational dependent-marking is also commonly achieved by **adpositions**. As Zwicky (1992: 370) puts it, 'anything you can do with cases you can also do with adpositions, and vice versa'. There is no widely accepted cover term

[1] In German, such an abbreviation is virtually obligatory: 'nominative case' is *Nominativ* (not *?*Nominativkasus*), 'genitive case' is *Genitiv* (not *?*Genitivkasus*), and so on.

[2] Indeclinable words may well exhibit all the case values of corresponding declinable words. For instance, Russian indeclinables like *taksi* 'taxi' can be used as nominative, genitive, dative, etc. without any restriction. The case value has to be inferred from the context, e.g. modifiers agreeing in case.

[3] As Mel'čuk (1986: 36) observes, agreement case (found especially in Indo-European languages, but occasionally also elsewhere) does not fall under this definition and should be considered a different category-system.

for cases and adpositions, but the terms **flag** and **relator** have sometimes been used as terms which are neutral with respect to the case/adposition distinction. In practice, we find considerable overlap between adpositions and case inflection. Normally adpositions are considered to be separate words, whereas case inflections are thought to be expressed by morphological means. But these means are typically invariable affixes, and such affixes may look much like short (and perhaps cliticized) words. A range of widely applicable criteria for distinguishing clitics and affixes have been discussed in the literature (Zwicky and Pullum 1983, Haspelmath 2002: §8.3), but there are many cases where these criteria do not yield clear-cut results or are not applicable. Thus, linguists will have to live with some indeterminacy in this area.

It is not uncommon to find dependent-marking elements that are written as one word together with their hosts described as 'postpositions', and vice versa separately spelled elements are often described in terms of 'case'. While not rare, such notational and terminological practice is not standard. Some linguists prefer seemingly neutral terms such as 'case particle' (a term often used in Japanese linguistics for elements like *ga* 'nominative', *o* 'accusative', *ni* 'dative') or 'case marker', but if *case* is defined as an inflectional category-system, the term 'case particle' is contradictory (because particles are by definition words, not inflectional elements), and 'case marker' would have to mean 'inflectional case exponent'.

But unfortunately, the term *case* does not always mean 'inflectional category-system expressing dependency relations'. It can also refer to these relations themselves, following Fillmore's (1968) terminological choice. Fillmore's intention was to highlight the importance of abstract semantic roles for languages like English that have (almost) no case distinctions. Rather than introducing a new term, he used the term *case*, familiar from case-inflecting languages, where cases primarily serve to express semantic roles (although few if any languages show a one-to-one mapping between cases and semantic roles). To distinguish the Fillmorean cases from the usual case concept, they have sometimes been called *deep cases* (because Fillmore claimed that they were universally present at 'deep structure'), *case roles*, or *case relations*, but it seems simpler and less confusing to call them *semantic roles*, a framework-neutral term that by now has wide currency (although it did not exist in the mid-1960s).

Another extension of the term *case* is due to Chomsky (1981), who used *Case* (often capitalized, to distinguish it from inflectional case) for an abstract property of noun phrases that in his Government-Binding theory licenses their occurrence (also called *abstract Case*). Case in this sense need not be overt, i.e. even isolating languages like Vietnamese require Case on all their NPs. But when a language has inflectional case, this is thought to be a manifestation of abstract Case. Thus, Case has a sense very similar to *grammatical relation* (a term that is generally avoided in Chomskyan syntax). Unlike Fillmore's extended case concept, Chomsky's extended Case concept has not been used beyond the framework in which it originated.

33.3 Kinds of cases

33.3.1 Grammatical cases vs. concrete cases

A distinction is often made between more abstract cases expressing core syntactic relations such as subject and object, and more concrete cases that express various specific semantic roles, especially spatial relationships (cf. Blake 1994: 32–4). Different term pairs have been used for these two classes of cases:

(1)	grammatical cases	semantic cases	e.g. Blake (1994: 32)
	relational cases	adverbial cases	e.g. Bergsland (1997)
	grammatical cases	concrete cases	e.g. Jespersen (1924: 185)
	core cases	peripheral cases	e.g. Blake (1994: 34)
	abstract cases	concrete cases	e.g. Lyons (1968: 295)

The distinction is made in different ways by different authors and for different languages, but the basic intuition behind it seems to be the same.

33.3.2 Structural and inherent Case

In Chomskyan syntax, the distinction between *structural Case* and *inherent Case* is somewhat similar to the distinctions of §33.3.1. Structural Case is case that is assigned in a particular structural configuration (e.g. accusative in the complement position of VP, nominative in the specifier of INFL, in the framework of the 1980s), while the assignment of inherent case is tied to a particular semantic role ('theta-role'), or to lexical properties of the governing head (e.g. dative case assigned by the German verb *helfen* 'help').[4] The latter kind of case is also known as *quirky case*, especially when the NP bearing the lexically determined case can be regarded as the subject (as happens famously in Icelandic).

33.3.3 Oblique cases

In another spatial metaphor going back to the ancient Greeks, the term *oblique* is used for all cases apart from the basic case (in Greek and Latin, the nominative).[5] This term is useful especially when the oblique cases share a formal property that is not shared by the nominative (for instance, Latin nouns sometimes have a different stem for the oblique cases, as in nominative *homo* 'human being', accusative *homin-em*, dative *homin-i*, etc.). In languages with a two-term case system, the term *oblique*

[4] When inherent case is lexically determined, it is also called *lexical case* (see Woolford 2006 for the distinction between inherent and lexical case in the Chomskyan framework).

[5] The nominative is sometimes called *casus rectus* 'direct case' in the older tradition. (The vocative is also considered a non-oblique case.)

(also *general oblique*) is often used as a case label for the single non-basic case, and the basic case is then called *direct case* (e.g. in Iranian and some Uto-Aztecan languages).[6]

33.4 Labelling cases

33.4.1 Alternatives to labelling

Labelling individual cases, i.e. referring to them by case labels such as *nominative*, *dative*, *instrumental* has proved very practical in linguistics, and such case labels will take up the remainder of the discussion of this chapter. However, there are at least two alternatives: First, cases can be referred to by the shape of their (primary) exponents. Some grammars actually do this and avoid labelling the cases, e.g. Bromley's (1981) grammar of Lower Grand Valley Dani (see 1981: 78), Seiler's (1977) grammar of Cahuilla (see 1977: 81–3), and Gordon's (1986) grammar of Maricopa (see 1986: 36).

Second, cases can be referred to by numbers. This is done, for example, in the traditional primary-school terminology for German cases (*1. Fall* 'nominative', *2. Fall* 'genitive', *3. Fall* 'dative', and *4. Fall* 'accusative'), and case numbering is occasionally used in modern descriptive grammars (e.g. in Tamura's (2000) grammar of Ainu). Numbering is particularly effective if the cases are arranged in a consistent, conventional order, and indeed, Plank (1991b) observes that the Western tradition since the Stoics used to put a lot of weight on the order in which cases are presented.

But in general, linguists opt for descriptive case labels, just as with most other inflectional category-systems (an exception being person, where numbering of the individual categories has been prevalent since antiquity).

33.4.2 Non-case uses of case labels

Not uncommonly, the descriptive labels that were created for cases are also used to label adpositions, e.g. by Guillaume (2004: ch. 14) in his grammar of Cavineña, and by Kießling (1994: 192–3) in his grammar of Burunge. This is perfectly reasonable, because adpositions function in much the same way as cases in languages, the main difference being that they are analytic means of expression. Thus, talking about the English 'dative preposition *to*' or the French 'genitive preposition *de*' is completely

[6] More recently, the term *oblique* has also come to be used for peripheral grammatical relations (expressed by peripheral cases or adpositions), following the terminology of Relational Grammar.

unproblematic (just as we can have both future-tense affixes and analytic future-tense auxiliaries).

Case labels can also be used to label semantic roles. Fillmore (1968) used some of the labels from the Latin tradition for his semantic roles, e.g. dative and locative. Especially the labels of the peripheral cases are quite close to semantic roles, so they can be used for the semantic roles also when these are not expressed by cases. Thus, we can say that in some languages, 'the comitative relation' is expressed by a serial verb construction, or that 'an instrumental applicative' is used where other languages would have an instrumental case or adposition.

Finally, the abstract case labels can be used to describe the alignment of other phenomena apart from case. In an ergative construction, the intransitive S and the monotransitive P are typically treated in the same way with respect to case-marking, but they may also be treated in the same way with respect to indexing on the verb. We may then say that a verb has 'ergative' and 'absolutive' person–number affixes, even when the language has no case-marking at all (as is the case in the Mayan languages, for example). Likewise, for languages with accusative alignment of person–number affixes, speaking of 'nominative' and 'accusative' person-number markers is appropriate (though the less precise labels 'subject markers' and 'object markers' are more common).

33.4.3 Cases as language-particular categories

Like all other morphosyntactic categories, cases are language-particular entities.[7] This means that case labels are valid only for particular languages. When talking about Latin, an expression like 'the dative case' has to be interpreted as 'the Latin dative case'; it is a kind of proper name for a unique category. For convenience, similar case labels are used for different languages, so that we also talk about 'the dative case' in Turkish, for example. But the transfer of case labels from one language to another should not be understood as meaning that we are dealing with 'the same case'. It is not meaningful to talk about 'the dative case' as such, regardless of particular languages, and when comparing categories of two languages, it is senseless to use formulations such as 'The dative case has a broader range of uses in Latin than in Turkish'[8] (the correct version of this would be: 'The Latin dative case has a broader range of uses than the Turkish dative case').

[7] Some linguists seem to assume that many categories (or the underlying features defining them) such as noun, verb, singular, future, animate, are innate, so that language-particular categories are instantiations of these innate universal categories. I disagree with this view but have no space to discuss it here. In any event, innateness has not been explicitly claimed for cases in any recent prominent publications, as far as I am aware.

[8] This sentence makes about as much sense as the sentence 'San Cristóbal has more inhabitants in Mexico than in Venezuela' (talking about two different cities both called San Cristóbal).

That categories are language-particular entities has been widely recognized at least since Saussure and Boas, and American structuralists have sometimes drawn the conclusion from this that idiosyncratic, opaque category labels (e.g. numbers) should be used in order to avoid the impression that one language is described in terms of the categories of another language (such as Latin or English). This concern is well-founded, but the more recent consensus is that opaque category labels make using a description very cumbersome. Grammatical descriptions are far easier to understand and remember by human readers if transparent and familiar labels are used. To make clear that we are dealing with language-particular categories, some authors have advocated capitalization of language-particular category labels, in the manner of proper names (Comrie 1976a: 10, Bybee 1985, Croft 2001: 51), and many grammars now follow this proposal (e.g. Haspelmath 1993b, Maslova 2003). The above statement would thus read: 'The Latin Dative case has a broader range of uses than the Turkish Dative case.'

33.4.4 Case polysemy

Some cases have just a single identifiable meaning, and ideally their label should reflect this meaning. For example, Chantyal has a special case for expressing the standard of comparison, and it is appropriately labeled *comparative case* (Noonan 2003: 320). But very often, cases have a range of meanings, and they are best described as being polysemous. To give a simple example, the Turkish Dative expresses recipient (*kıral-a* 'to the king') and direction (*Trabzon-a* 'to Trabzon'), and the Latin Dative case expresses possessor (*Flavi-o est liber* 'Flavius has a book') and beneficiary (*vitae discimus* 'we learn for life'), among other functions.

To reflect the meanings of such cases faithfully in their labels, one could use multiple-term labels, e.g. *dative-allative* for Turkish, and *possessive-benefactive* for Latin. And indeed, such double names are very common in the literature (e.g. *allatif/datif* and *instrumental/locatif* for Tunumiisut Eskimo, Mennecier 1995: 252; *locative/allative* and *locative/illative* for Classical Tibetan, DeLancey 2003: 258). However, it should be kept in mind that case labels can never capture the full range of semantic and other properties of case. Case labels are primarily mnemonic devices, and they should reflect some important semantic properties of the case they designate. But the full range of uses has to be described separately anyway, so complex case labels are not really necessary. As long as readers are aware that cases are language-particular categories, they will not draw the wrong inferences from short case labels which fail to be fully descriptive of the cases they designate. Since many cases have a fair range of distinguishable meanings, fully descriptive case labels are not practical anyway (the Latin Dative case would have to be called at least 'dative-possessive-benefactive-experiential'). The best method for comparing the range of uses of cases across languages is the semantic-map

method (Haspelmath 2003; see Narrog and Ito 2007 for a recent application to cases).

Another way of dealing with case polysemy that is occasionally found is to split a case into several different labelled entities. For instance, Miller (2001: 157–8) uses several different names for the *-m* case of Jamul Tiipay ('instrumental, comitative, allative'), and glosses the case differently depending on the English equivalent. Since the different translations into other languages seem to be the main reason for these label choices, a better option would be to pick one of the labels as a mnemonic device (say, Comitative), and specify that the Jamul Tiipay Comitative can also be used in instrumental and allative senses.

33.5 Case labels: abstract cases

The term **nominative** is generally used for the S (single argument of intransitive clause) and A (most agent-like argument of the transitive clause), and in most languages this is also the (zero-coded) citation form of the noun (*nominativus*, the Latin rendering of Greek *onomastikḗ*, originally meant 'naming form'). In languages with ergative alignment of case-marking, the citation form of the noun almost always occurs both as S and as P (most patient-like argument of transitive clause), and this case is now mostly called **absolutive**.

The case of the P in accusative alignment is called **accusative**, and the case of the A in ergative alignment is called **ergative** (in fact, the alignment types were named after the cases). The terms *nominative* and *accusative* are very old, whereas the term *ergative* only became widespread in the first half of the twentieth century (starting with Caucasian linguistics; see Manaster-Ramer 1994, Butt 2006: 154–8), and the term *absolutive* only became widespread in the second half of the twentieth century (starting with Eskimo linguistics). However, especially in Caucasian linguistics, this case is still often called *nominative*, thus reflecting the etymological sense of the term (Mel'čuk 1988: 208, Blake 1994: 187, n. 4). For 'ergative', other terms are still used in some traditions (*relative case* in Eskimo linguistics, *narrative case* in Kartvelian linguistics). Earlier term pairs for 'ergative/absolutive', which are now obsolete, are *casus activus/casus passivus* (Jespersen 1924: 166) and *nominativus agentis/casus indifferens* (Pilhofer 1933: 44).

A problem arises in languages where the S and A are coded alike, but exceptionally the S/A case is not zero-coded, but is overtly coded ('marked nominative' constructions). In such languages, there is a tendency to call the (zero-marked) case of the P not 'accusative', but *absolut(iv)e case*, and to call the marked nominative not 'nominative', but *subject(ive) case*.

The labels **subjective case** and **objective case** are sometimes used for English instead of *nominative/accusative* (e.g. Quirk et al. 1985: 337), perhaps in order to highlight that the English (pronominal) case system is organized very differently from the Latin case system. Iggesen (2005b: 92) points out that in languages where the P case is also used for the recipient of ditransitive clauses, it cannot properly be called *accusative* and proposes to call it *objective*.

A general problem with the terms *nominative, accusative, absolutive, ergative* is that they are well-defined only for idealized systems without splits. But splits are very common – for instance, we find many languages with accusative marking only on definite and/or animate Ps. Some authors find it awkward to call the case of the zero-coded P in such languages nominative (e.g. Gorelova 2002: 163 calls the case in Manchu *casus indefinitus*). Similarly, in languages that have a subclass of nominals with overt ergative marking in A function (e.g. all third person nominals), and a subclass of nominals with overt accusative marking in P function (e.g. all personal pronouns), it is unclear what label should be used for the zero-coded case(s), *nominative* or *absolutive*. No general solution seems to be available, but since case labels are primarily menomonic devices, ad-hoc solutions are adequate as well.

Where the intransitive S is split into agent-like S_A (coded like A) and patient-like S_P (coded like P), the case for S_A/A is best called **agentive**, and the case for S_P/P is best called **patientive** (see Iggesen 2005b: 93), though such cases are not common. *Agentive case* is also used for the case of the demoted passive agent (e.g. in Poudel's (2006) grammar of Dhankute Tamang (p. 102)).

The term **genitive** for the case of the possessor is relatively unproblematic, and there seems to be no strong reason to rename it *possessive case* (as is done, for instance, in Charney's 1993 grammar of Comanche). However, it should be borne in mind that in typological studies, *genitive* is also often used to refer to the possessor in possessed NPs, regardless of whether it is expressed by a case (or adposition). Some languages have an overt marker on the possessum in possessive constructions, and calling such a marker 'genitive' as well is confusing (see e.g. Frajzyngier 2002: 50 for Hdi). (A label such as *antigenitive* for such markers seems more appropriate, but since they are not cases, I will not say more about them here.)

The term **dative** for the case of the indirect object (i.e. the recipient argument that is marked differently from the monotransitive P; see Dryer 1986, Haspelmath 2005a) presents no difficulties. Since the recipient is often coded in the same way as a direction, linguists sometimes vacillate between *dative* and *allative*, but both are equally good terms for a case with both uses. When the recipient (or R) of a ditransitive construction is coded like the monotransitive P, we are dealing with a 'primary object', and following Haspelmath (2005a) the corresponding case could be called **primative** (or *objective case*, as suggested by Iggesen 2005b). Where there is a special case for the ditransitive theme (the 'secondary object'), this could be

978–0–19–920647–6 34-Spencer-ch33 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 514 of 517 May 16, 2008 18:59

called **secundative** (though primative and secundative cases are rare; secundativity is mostly found in indexing patterns).

Somewhat more concrete cases related to datives are **benefactive** (used e.g. in Basque, Hualde and Ortiz de Urbina 2003: 183) and **destinative** ('intended for'), e.g. in Kâte (Pilhofer 1933: 44), Basque, and Udihe (Nikolaeva and Tolskaya 2001: 126). A special case for experiencers is called **affective** in Daghestanian linguistics (e.g. in Godoberi, Kibrik 1996: 16).

Some languages have special cases for predicate nominals, called **essive** (e.g. Finnish), or **predicative** (e.g. Yukaghir, Maslova 2003: 91), and also for predicate nominals of verbs of change ('become something, turn into something'), called **mutative** (in Ainu, Refsing 1986) or **translative** (in Khanty).

The term **partitive** is used for a very special case (having to do with partial affectedness of an object argument) in Finnish, and for a similar case in Basque, but has apparently not been found useful for many other languages. The term is also often used for certain case/adposition uses (e.g. 'five of my friends'), but no language seems to have a special case just for this partitive function.

Finally, the *vocative* form of the noun is often considered a case, following the Greek and Latin tradition, but it clearly does not fall under the standard definition of case, and there are only very few languages in which the vocative shows similarities with cases (such as triggering agreement on modifiers; see Daniel and Spencer, Chapter 43).

33.6 Case labels: concrete nonspatial cases

What is called here 'concrete nonspatial case' is a very heterogeneous group that is set up here only for expository convenience. The most important case labels in this group are **instrumental** ('with, using') and **comitative** ('together with'). Since none of the older Indo-European languages has a comitative, this label was slow to catch on, and competing terms are still sometimes used: *associative* (e.g. for Sahaptin, Rude 1997; Maricopa, Gordon 1986: 42; and Abun, Berry and Berry 1999), *sociative* (especially for South Asian languages, e.g. for Dhivehi, Cain and Gair 2000; Korku, Nagaraja 1999: 49), *accompanitive* (Sye, Crowley 1998a: 293). Another case label related to the comitative is **proprietive** 'with, having', common especially in Australian languages. The *ornative* case of Dumi ('endowed with, equipped with', van Driem 1993: 76) is similar.

The negative counterparts of these 'with' cases are 'without' cases, variously called **abessive** (e.g. in Uralic languages), *caritive* (e.g. in Kâte, Pilhofer 1933: 45),

978–0–19–920647–6 34-Spencer-ch33 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 515 of 517 May 16, 2008 18:59

privative (e.g. in Chukchi), *anticomitative* (also in Chukchi, Kämpfe and Volodin 1995: 30), or *deprivative* (e.g. in Gooniyandi).

Some languages have a **comparative** case for the standard of comparison of inequality (e.g. Chantyal, Noonan 2003: 320; Dumi, van Driem 1993: 78), and quite a few have a **similative** case to express 'like'. Alternative labels for the latter are *simulative* (Eskimo, Mennecier 1995: 460), *similitive* (Sye, Crowley 1998a: 220; Sedang, Smith 1979: 125), *similaritive* (Yamphu, Rutgers 1998: 75), *semblative* (Australian languages), and *equative* (Eskimo, Holst 2005: 94).

Various causal relations are also sometimes expressed by cases. Some authors mention a **causal** case (e.g. Hosokawa 1991: 275 for Yawuru), and the *motivative* of Basque ('because of', Saltarelli 1988: 300), as well as the *consequential* of Kayardild (Evans 1995a) are equivalent terms. Many Australian languages are described as having an **aversive** case ('for fear of', 'to avoid'), also called *negative causative* (e.g. Yallop 1977: 75 for Alyawarra). Evans (1995a) distinguishes a *utilitive* case in Kayardild ('used for').

Some languages have a case whose only use is with (certain) adpositions. Since such cases have no meaning, they can be called simply *prepositive/prepositional* (the best-known example is Russian).

33.7 Case labels: Spatial cases

Cases expressing spatial relations can be grouped into four broad directional classes: cases expressing location ('at'), goal ('to'), source ('from'), and path ('through, along'). The basic terms for the first three are **locative**, **allative**, and **ablative**. Alternative synonym terms for 'allative' are *directional* (e.g. in Basque, Hualde and Ortiz de Urbina 2003: 186), *directive* (e.g. in Sumerian, Balke 1999: 121), and *lative* (e.g. in Udihe, Nikolaeva and Tolskaya 2001: 124).

For cases denoting a path, there is little terminological agreement. Perhaps the most common label is **perlative** ('along, through'), e.g. in Wambaya, Nordlinger (1998a: 91); Lavukaleve, Terrill (2003: 64); Cavineña, Guillaume (2004: 550). Synonyms are *prosecutive* (e.g. in West Greenlandic Eskimo, Fortescue 1984: 206), *prolative* (e.g. in Udihe, Nikolaeva and Tolskaya 2001: 125), *traversal* (e.g. in Ainu, Refsing 1986: 168), *translative* (e.g. in Archi, Kibrik et al. 1977: 2.59), *vialis* (the older literature on Eskimo), and *mediative* (e.g. Belhare, Bickel 2003: 549).

For a movement that goes all the way to its endpoint, the case label **terminative** has been used (e.g. in Basque, Hualde and Ortiz de Urbina 2003: 186), or its synonym *limitative* (e.g. Ika, Frank 1990: 38).

For a movement that goes only in the direction of (or 'toward') its goal, the term **orientative** has been used (e.g. for Kham, Watters 2002: 690), or its synonyms

adversive (for Kâte, Pilhofer 1933: 46–7), and *versative* (for Tsez, Comrie and Polinsky 1998: 104).

In addition to these directional distinctions, cases sometimes also express orientational contrasts such as 'in' vs. 'on' vs. 'at'. In Hungarian, for instance, there is a contrast between the **inessive** (e.g. *a táská-ban* 'in the bag'), the **superessive** (e.g. *a ház-on* 'on the house'), and the **adessive** (*a hajó-nál* 'at/near the ship'). This three-way orientational contrast can be combined with the directional contrast between location, goal, and source, yielding nine spatial cases (see also Creissels, Chapter 42):

(2)

	'in'	'on'	'at'
location	inessive (*-ban*)	superessive (*-on*)	adessive (*-nál*)
goal	illative (*-ba*)	sublative (*-ra*)	allative (*-hoz*)
source	elative (*-ból*)	delative (*-ról*)	ablative (*-tól*)

The case labels used for Hungarian (and similarly for other Uralic languages) are mostly based on Latin verbs such as *il-lat-* 'carry in', *e-lat-* 'carry out', *sub-lat-* 'carry up', *de-lat-* 'carry off'. A more transparent system for labelling the Hungarian cases would be the system in (3), where the three directions are consistently expressed by the prefixes *in-*, *super-*, and *ad-*, while the three orientations are consistently expressed by the stems *-essive*, *-allative*, and *-ablative*.

(3)

	'in'	'on'	'at'
location	in-essive	super-essive	ad-essive
goal	in-allative	super-allative	ad-allative
source	in-ablative	super-ablative	ad-ablative

In some Daghestanian languages, which tend to have very transparent systems of spatial cases expressing direction and orientation by different markers, such a system is used. For Lezgian, Haspelmath (1993b: 74) uses the following labels (he used *-directive* and *-elative* instead of *-allative* and *-ablative*):

(4)

	location	goal	source
'in'	in-essive (*-a/-e*)	(dative)	in-elative (*-aj/-äj*)
'on'	super-essive (*-l*)	super-directive (*-l-di*)	super-elative (*-l-aj*)
'at'	ad-essive (*-w*)	ad-directive (*-w-di*)	ad-elative (*-w-aj*)
'behind'	post-essive (*-q^h*)	post-directive (*-q^h-di*)	post-elative (*-q^h-aj*)
'under'	sub-essive (*-k*)	sub-directive (*-k-di*)	sub-elative (*-k-aj*)

Other Daghestanian languages are much richer and have additional directions such as **translative** 'through, along' and **versative** 'toward', as well as further orientations such as 'on (horizontal)' vs. 'on (vertical)' (*cont-*), 'near' (*apud-*), 'among' (*inter-*) (e.g. Kibrik 1996: 17, van den Berg 1995: 45; see also Daniel and Ganenkov, Chapter 46). By combining directions and orientations with each other, and adding further markers for deictic distinctions, over 100 spatial cases can be distinguished.

But Comrie and Polinsky (1998) and Comrie (1999) point out that these 'cases' are not single inflectional categories, but combinations of categories from at least two different inflectional category-systems. Already Kibrik et al. (1977: 2.51) had set up a separate inflectional category-system **localization** for the different orientation markers, which combine with different spatial case-markers. On this view, a label such as *super-elative* would not stand for a single case, but for a localization–case combination (analogous to labels such as *past perfect* or *pluperfect*, which stands for a particular tense–aspect combination).

CHAPTER 34

CASE POLYSEMY

ANDREJ MALCHUKOV
HEIKO NARROG

34.1 Introduction: polysemy and the semantic map approach

It is probably true that cases which are restricted to one specific meaning are rarer than cases subsuming several meanings. The origins of these polysemies or syncretisms vary: some are conditioned by phonetic factors, especially pervasive in the situation when a case system collapses (see Barðdal and Kulikov, Chapter 30), others have a transparent semantic basis. An example of the former type is the dative–ablative–instrumental syncretism in Greek, which was conditioned by a merger of three cases into one form (Luraghi 2003). The latter type may be exemplified by the dative–allative polysemy familiar from many European languages (cf. English preposition *to*: *gave to a friend* vs. *came to a station*). While the distinction between the two types may be hard to draw in cases where both phonetic and semantic factors contribute to polysemies (see Baerman, Chapter 14, and Iggesen 2005b for discussion[1]), cross-linguistic comparison is usually helpful. Insofar as the form of markers is idiosyncratic across languages, the syncretism of the former type will be

[1] Baerman, Chapter 14, notes that even syncretism conditioned by phonetic factors may give rise to analogical extension across declension paradigms, a process presumably aided by some functional similarities between the forms conceived by the speakers.

cross-linguistically rare, while the semantically driven polysemies will be recurrent across languages.

A semantic map is a method of capturing and accounting for regular polysemies of linguistic forms. It starts from the iconicity assumption that recurrent similarities in form reflect regularities in meaning. This idea has been looming in functional linguistics,[2] but it was first explicitly formulated in the work of Anderson who proposed a semantic map for temporal and evidential domains (Anderson 1982; 1985). Since then the semantic map approach has been applied to a number of different domains including case; here the work by Haspelmath (2003) should be particularly acknowledged.[3] The basic methodology behind the semantic map approach as fleshed out by Croft (2001) and Haspelmath (2003) relies on the establishment of cross-linguistic regular polysemies, which on iconicity assumptions reflect conceptual similarities. The functions covered by the same form in (some) languages are considered to be semantically close, which is represented by putting them adjacently in the semantic space. Although the semantic map is established empirically through the study of polysemy patterns across languages, the established semantic connections are conceived as universal, and give rise to predictions concerning possible and impossible[4] polysemy patterns across languages.

This methodology can be explained on the basis of a simple map, including directional (*go to London*), recipient (*give a book to John*), and beneficiary functions (*cook meal for a child*). Comparing English data to German or Russian we note that while English uses the same marker for the former two functions and marks beneficiaries distinctly, Russian and German code both recipients and beneficiaries by dative case, using directional prepositions (Russian *v*. German *nach*) for the first function.

[2] Haspelmath (2003: 216) refers in particular to Haiman's (1985 *et passim*) work on iconicity. Also other work in the cognitive-functional tradition such as 'semantic networks' in cognitive grammar (see Luraghi, Chapter 9), or the 'chain polysemy model' of Wierzbicka (1980), where selective semantic affinities between individual functions rather than a general meaning is postulated, have similarities to the semantic map approach. Yet, these models are not conceived as giving rise to (implicational) universals (Haspelmath 2003).

[3] An important predecessor to the semantic map approach to case polysemies is an article by Frans Plank dealing with case syncretism in different Indo-European languages (Plank 1991b). Plank proposes to capture case syncretisms through graphs reminiscent of semantic maps. Interestingly, the author concludes that it is impossible to arrive at one single configuration of case homonymies that would hold for all languages, although he notes some general tendencies. This may be due to the fact that the author does not distinguish consistently between the cases of systematic vs. accidental case syncretisms (see Baerman, Chapter 14, on this distinction). It should be noted, however, in this context that some current approaches to case syncretism (building on the work by Jakobson and Bierwisch) do try to account for all patterns of syncretism in terms of abstract (syntactic) features and feature underspecification (see Müller et al. 2004 for a collection of papers representative of this approach).

[4] Or rather: probable and improbable, since a semantic map in individual cases may be violated due to interference of other factors.

Table 34.1. Ngiyambaa (adapted from Donaldson 1980: 86)

Functions	Nouns	3rd person (bound) pronouns	1st, 2nd person (free) pronouns
Instrument	ERG	–	–
A(gent)		ERG	NOM
S (intransitive subject)	ABS	ABS	
P, R (direct and indirect object)			OBL
Possessor	DAT	GEN	
Goal (allative)		–	OBL+LOC
Locative	LOC	–	
Ablative	CIRC	–	OBL+CIRC

This suggests a semantic configuration where Recipient is intermediate between Goal and Beneficiary, with which Recipient shares also functional similarities:

Goal – Recipient – Beneficiary

This configuration is assumed to be universal; in fact it is a subpart of a semantic map of dative functions proposed by Haspelmath (see below).

It should be noted that although semantic maps were established on the basis of cross-linguistic patterns, it is possible to come up with a plausible map for a single language, provided that this language has several markers(for example, German has no less than five distinct markers for the instrument function all of which are polysemous in a different way; Narrog, Chapter 40), or that different nominals display different types of polysemies. Australian languages are representative of this type of case asymmetry (cf. Iggesen, Chapter 16). This is illustrated below for the case of Ngiyambaa where different types of nominals show different polysemy patterns (cf. Dench, Chapter 52, for a similar chart for another Australian language, Nyamal). Table 34.1 (adapted from Donaldson 1980: 86) shows the use of seven cases across different declension types: ERG (ergative–instrumental), ABS (absolutive), DAT (dative), LOC (locative), CIRC (ablative–circumstantial), GEN (genitive), and OBL (oblique), the latter two occurring only with pronouns. Apart from the difference between nouns and pronouns in alignment types (ergative with nouns, accusative with pronouns), this table demonstrates a number of syncretisms involving oblique cases: ergative–instrumental, dative–genitive, locative–dative. The latter syncretisms seem non-accidental as well,[5] as they are well attested cross-linguistically (see below).

[5] See though Baerman, Chapter 14, for a somewhat different view.

34.2 Core arguments: polysemy patterns and syntactic alignment

Polysemy patterns involving encoding of core arguments of intransitive (S) and transitive (A, P) is a traditional domain of investigation of alignment typology (cf. Bickel and Nichols, Chapter 20). One of the early applications of this approach was Fillmore (1968), where we find alignment schemes reminiscent of semantic maps, visualizing identical marking of core arguments in accusative alignment (A=S≠O), and ergative alignment (A≠S=O). This approach has become standard in typology from the 1970s, due to work by Comrie, Dixon, Kibrik, and others, and is still maintained in the modern typological literature (see e.g. Song 2001; cf. however, Bickel and Nichols, Chapter 20, for a critical assessment). Later on, alignment typology has been extended by Dryer (1986), Haspelmath (2004), and Siewierska (2004) into the domain of ditransitive constructions. In the latter domain the main distinction is between languages distinguishing between direct and indirect objects (P=T≠R), and languages distinguishing between primary vs. secondary objects (P=R≠T). In Haspelmath's terms, the former languages have 'indirective' alignment, and the latter 'secundative' (see Bickel and Nichols, Chapter 20; Siewierska and Bakker, Chapter 19; Kittilä and Malchukov, Chapter 36, for further discussion of alignment patterns in ditransitive clauses and exemplification).

It should be noted that alignment patterns have not been conceived originally as semantic maps, but rather as possible groupings of core arguments on the basis of their encoding. This approach is perhaps most clearly represented by Kibrik (1979; 1985; cf. also Plank 1985), who discusses different alignment patterns from a multi-factorial perspective. While he admits that proper encoding of semantic roles is an important factor (which, for example, account for the marginality of the alignment pattern grouping together A and P arguments to the exclusion of S: S≠A=O), he also acknowledges other factors such as economy or distinguishability of arguments. Thus, languages with a neutral alignment (A=S=O) do not distinguish between arguments, while tripartite systems (A≠S≠O) are uneconomical (cf. Song 2001). Yet, more recently the alignment schemes have been explicitly reformulated in terms of semantic maps. Thus, Croft (2001) proposed the semantic map in Figure 34.1 for the encoding of core arguments.

Given standard assumptions about well-formedness of semantic maps, this approach correctly predicts the marginality of alignment types which would display discontinuous segments on the map (such as the semantically anomalous S≠A=O pattern). It also conveniently captures the variation of alignment patterns across languages. Thus the following two maps (Figures 34.2–3) represent two of the possible alignment types complying with the semantic map: the accusative 'indirective' language (such as German, see (1)), and an ergative language with a 'secundative' alignment, such as Eskimo (see (2)).

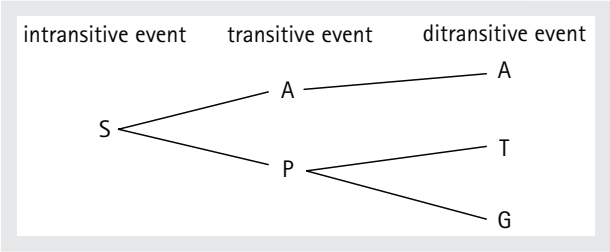


Figure 34.1. Croft's conceptual space for core arguments (participant roles)

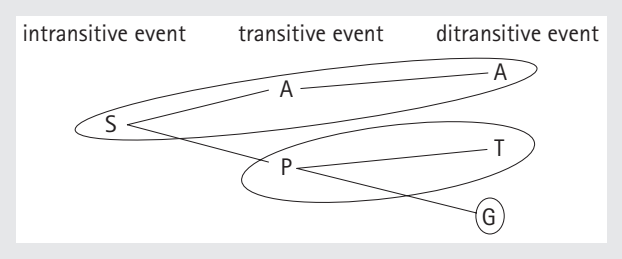


Figure 34.2. Alignment maps: German

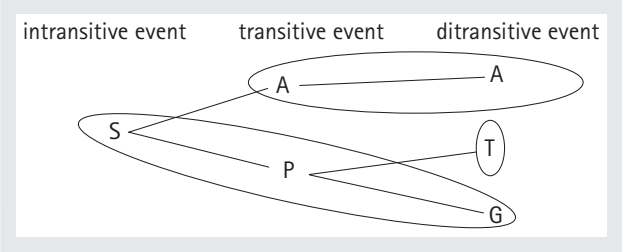


Figure 34.3. Alignment maps: Eskimo

(1) *Ich gab ihm ein Buch*
'I(NOM) gave him(DAT) a book'

(2) Eskimo (West Greenlandic; Fortescue 1984: 88)
(Uuma) Niisi aningaasa-nik tuni-vaa
(that.ERG) Niisi money-INSTR.PL give-IND.3s-> 3s
'(He) gave Nisi money'

Similar maps can be drawn to represent other major alignment types, which are largely consistent with the map insofar as they do not involve contiguity violations[6]

[6] For some exceptional patterns like the double oblique pattern attested in some Iranian languages, see Arkadiev, Chapter 47.

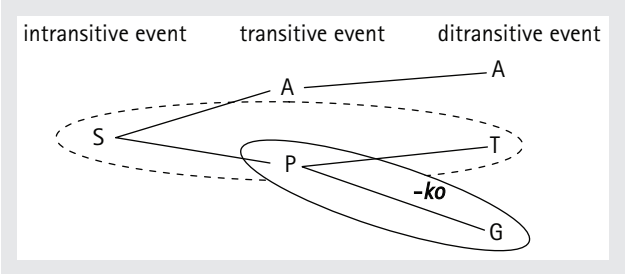


Figure 34.4. Hindi: Alignment pattern complicated by DOM

(cf. Dryer 2007 for similar diagrams, which however are not conceived as semantic maps).[7]

One of the phenomena which are not captured by the alignment map in Figure 34.1 is differential case marking of arguments, where different types of nominals select for different cases (see Malchukov and de Swart, Chapter 22). Yet it is perfectly possible to augment maps in a way that will capture these more complex patterns. Thus the situation in Hindi where animate NPs should be marked by accusative/dative case-marker *-ko*, while inanimates are marked only if definite, can be represented by Figure 34.4 (the dashed line represents absence of overt case marking).

This pattern holds for the encoding of arguments in imperfective tenses; the picture gets more complicated once ergative patterns in the perfective tenses are taken into account as well. The situation gets still more complex once ditransitive constructions with animate T are considered. Here again, several distinct patterns can be discerned, depending on whether Differential Object Marking is extended to ditransitive constructions with animate themes or suspended (see Kittilä and Malchukov, Chapter 36, section 1.3).

Note that Croft's map is unconventional in the sense that it differentiates between subjects of monotransitives and ditransitives, which do not differ ostensibly in role properties and are not differentiated in the absolute majority of languages.[8] Also the status of S as a single (macro)role is questionable. Evidence from split intransitive languages suggests that it might be useful to split S into agentive and patientive S roles (cf. Dixon 1994).[9] Generally, uniform marking of intransitive subjects across

[7] It should be also noted that the proposed maps are designed to capture major constructions and minor patterns are disregarded in these representations (i.e. the canonical transitive clauses rather than, say, constructions with dative experiencers are taken as representative of a bi-valent verb pattern). For a different approach see Bickel and Nichols, Chapter 20.

[8] See though Bickel and Nichols, Chapter 20 for some exceptions, which interestingly do not conform to the map, insofar as ditransitive A is unmarked for the ergative case unlike the A of monotransitives. See also Wichmann, Chapter 56, on the 'pegative' case in Tlapanec.

[9] Cf. also König, Chapter 35, for the issue of differentiated marking of S.

languages does not have a semantic basis, but is due to different factors such as economy and distinguishability (cf. Kibrik 1985). This highlights a question that application of semantic maps to core arguments is not unproblematic, since role properties are not the only motivating factor involved (see Malchukov and de Swart, Chapter 22, for discussion of indexing vs. differentiating functions of case marking). Yet, the semantic map approach is appealing as it constrains alignment patterns cross-linguistically, (correctly) predicting that contiguity violations are cross-linguistically infrequent. Naturally, this map may be extended beyond core arguments to obliques in order to represent the fact that R arguments frequently stem from directional markers, while dedicated T-markers as those in Eskimo frequently involve instrumental markers (see below).

34.3 Grammaticalization research; from oblique to core marking

In the preceding sections, research on the polysemy of case markers and its representation in maps was discussed from a synchronic perspective. However, if a specific linguistic form is polysemous, the possibility exists that one function or meaning is original in a historical sense, and that other functions and meanings are derived from it. The idea that polysemy is the product of historical change was espoused by Breal in the nineteenth century, and is also firmly supported in contemporary research on polysemy and meaning change (e.g. Traugott and Dasher 2002, Blank 2004). In addition, it is often assumed that semantic changes leading to the development of a new meaning for a specific linguistic item may be restricted to one direction. That is, while an item with meaning A can acquire meaning B, the reverse may not be the case. In the area of case functions, for example, it has often been claimed that a benefactive can assume dative function but not vice versa (cf. e.g. Lehmann 1995: 109f; Heine and Kuteva 2002: 54). If such unidirectionality between two or more meanings holds, historical information on the relationship between grammatical meanings and functions can add a new diachronic dimension to the semantic map. With the addition of this dimension, one can so to speak 'dynamicize' the map.

On these premises, a number of hypotheses on case polysemy from a diachronic perspective have been put forward in the framework of grammaticalization theory. The earliest, and up to now still most concrete and far-ranging proposal comes from Lehmann (1995: 109–12), who was primarily concerned with the development of core case marking, with the dative and the ergative as focal points. Lehmann claimed that there are two major sources for datives, namely directionals and

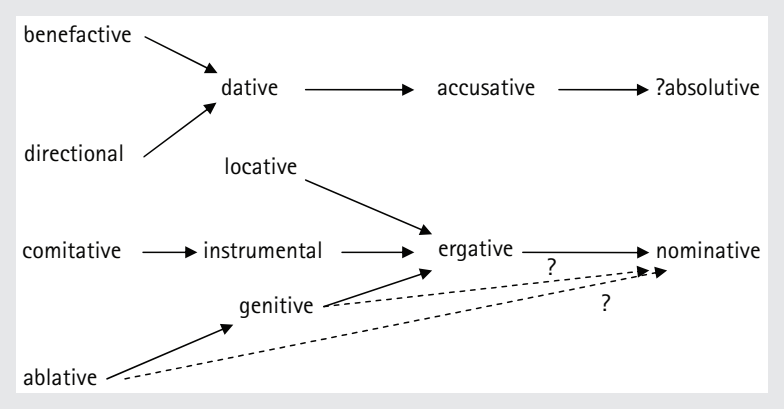


Figure 34.5. Grammaticalization channels between case markers according to Lehmann (1995: 112)

benefactives. A dative marker may further develop into an accusative marker. An ergative marker, according to Lehmann, can have even more sources, namely dative, locative, instrumental, or genitive marking, and can turn itself into a nominative marker. Nominative marking may have as its source also genitive and ablative marking.[10] Furthermore, on the periphery of the model, shifts from ablative to genitive and from comitative to instrumental are posited. These 'grammaticalization channels' are represented in a map in Figure 34.5 (the absolutive is marked by a question mark because, according to Lehmann, its development out of an accusative is not actually attested).[11] Note that this is a map of grammaticalization, but not necessarily a map of case polysemy. In some cases, Lehmann seems to assume that a marker with a specific case function 'develops into' a marker with a different function (e.g. dative > accusative; p. 110), and in other cases that it retains its old function (e.g. instrumental > ergative; ibid.). However, it is easy to reconceptualize the grammaticalization channels as connections on a dynamicized map of case polysemy. Recently, it has been a common assumption in grammaticalization theory that older meanings and functions are often retained when new meanings and functions develop, and not immediately disposed of (cf. Traugott and Dasher 2002: 11f.).

Interestingly, Lehmann assumes a parallel development of meaning and form. That is, in every single language, the continuum of functions is supposed to

[10] This seems to be less frequent (for the obvious reasons that nominatives are usually unmarked), but elsewhere (Lehmann 1995: 118) Lehmann mentions Burmese subject marker *-ka* of ablative origin, and Japanese *ga* of genitive origin).

[11] It should be noted that Lehmann's scheme, which remained unchanged in the 2002 second edition of his *Thoughts on Grammaticalization*, may not be complete or exhaustive. Creissels (p.c.) made us aware of the paths directional > locative, directional > genitive, and benefactive > genitive, which are documented in a number of languages.

correspond to a continuum of less to more grammaticalized forms from the left to the right, that is, for example, coverbs or relational nouns marking function on the left of the map to case markers or zero on the right (Lehmann 1995: 112).

Lehmann's hypotheses were based on historical observations in a variety of languages of which the author happened to have knowledge. Heine et al. (1991: 156–60), in contrast, based their work on conceptual considerations rather than historical data when they posited the following general chain of case functions:

(3) Chain of increasing grammaticalization of case functions

ablative		agent		purpose		time		condition		manner
allative	>	comitative	>	instrument	>		>	cause	>	
locative		benefactive		dative						
path				possessive						

(Heine et al. 1991: 159)

The functions further on the right are assumed to be more grammaticalized than the functions further on the left, the relationship being unidirectional. That is, functions on the left can be assumed to grammaticalize into functions on the right. Between functions that are listed together in one column, no difference with respect to grammaticalization is claimed. Again, this cline of grammaticalization does not directly signify case polysemy, but it implies a directionality in cases where case polysemy holds.

More recently, Heine, in Heine (2003: 595) and together with Kuteva in Heine and Kuteva (2002), proposed the following concrete paths of developments between case functions if a specific case function is taken as the source. The directionality was determined through reference to a wide range of work, including extant literature on grammaticalization, historical studies, and internal reconstruction from contemporary language descriptions.

(4) Paths of grammaticalization for case functions according to Heine and Kuteva (2002)

Ablative	>	agent, material, comparative, partitive, possessive
Allative	>	dative, patient, purpose, temporal, benefactive
Benefactive	>	dative, possessive, purpose
Comitative	>	agent, NP-and, instrument, manner, temporal
Dative	>	comparative, patient, possessive
Instrument	>	ergative, manner
Locative	>	agent, cause, comparative, possessive, temporal
Temporal	>	cause

The directionalities in the newer work appear to fall in line with the directionalities in Heine et al. (1991), except for the ergative, which was previously not included. Again, no polysemy is described here directly, but if polysemy occurs, it is implied that the direction of semantic extension conforms with the directionalities hypothesized here (see Heine, Chapter 29, for further discussion).

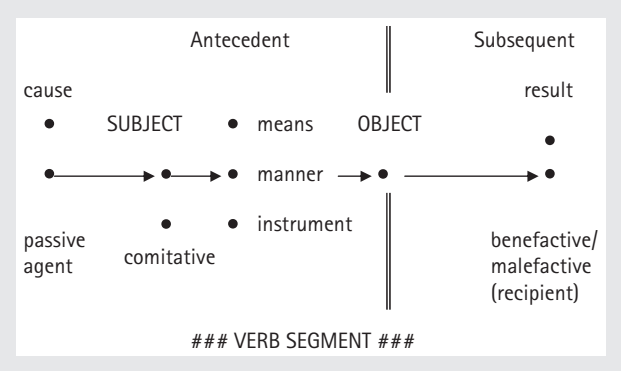


Figure 34.6. The causal chain model (Croft 1991: 185)

34.4 Uni-dimensional maps: Croft's (1991) causal chain model

A unique approach to case polysemy is taken by Croft in his causal chain model. In this model, 'the central concept around which the meanings (uses) of surface case markings are organized is the position of the participant in the verbal segment' (Croft 1991: 184). The verbal segment is characterized by a linear order of 'subject < object < oblique' which at the same time constitutes a grammatical relations hierarchy. There is a distinct cut-off point, or 'dividing line' in this linear chain, namely at the position of the 'object'. Consequently, case markers cluster their polysemies (in terms of thematic roles) either to the left of this dividing line, or to its right, that is, either to 'antecedent roles' or to 'subsequent roles'. Meaning extension ('spread') only takes place between elements that are semantically contiguous in the causal chain. The scheme is represented in Figure 34.6.

A survey of oblique case marking in forty languages largely confirms the model (Croft 1991: 187f., 237–9). In the majority of cases, case syncretism occurs either on the left side of the dividing line, or on the right side, but not across it. Thus, one finds markers combining comitative, instrumental, and manner function, but usually not comitative, instrumental, and result.[12] Yet some other polysemies seem to be more difficult to capture with this model (e.g. between the cause and instrument), and others are predicted to hold but are actually unattested (ergative/comitative; cf. Luraghi 2003). Note that case syncretism is the major evidence adduced by Croft to support the causal chain model, so in a way, it would be circular to predict case

[12] It should be noted that the vertical dimension of the chain model seems also to be significant, since those functions which are adjacent on that dimension tend to cluster under the same form (e.g. means/manner/instrument), although this issue is not explicated. On this assumption, the chain model may be viewed as a (two-dimensional) branching map.

978–0–19–920647–6 35-Spencer-ch34 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 528 of 534 May 16, 2008 18:59

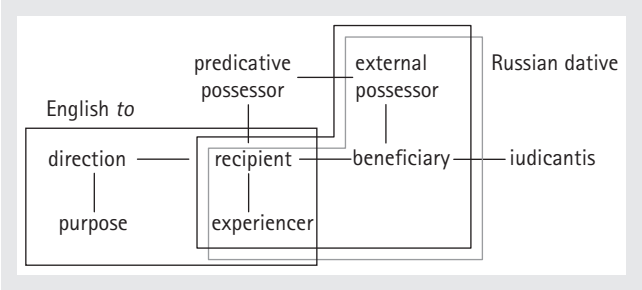


Figure 34.7. Semantic map of dative functions including the boundaries of English *to* and the Russian dative (cf. Haspelmath 2003: 213)

Au: Please provide citation for figure 34.7.

syncretism from it. Yet in spite of these limitations the model fares well in capturing the major polysemy patterns, and, being one-dimensional, supersedes in predictive power two-dimensional maps considered below.

34.5 Beyond core arguments: Datives and related functions

Haspelmath (2003; cf. Haspelmath 1999a) proposed the following map for the dative domain, including the following functions: a) direction (cf. *go to Leipzig*); b) recipient (cf. *gave an apple to me*); c) experiencer (*It seems outrageous to me*); d) purpose (*I left a party to get home early*), as well as others which are lacking for English *to* but available for French *à* or Russian dative: predicative possession (cf. French: *Ce chien est à moi* 'This dog is mine (lit. to me)'), dativus iudicantis ('judger's dative', as in German: *Das ist mir* (dat) *zu warm* 'This is too warm for me'), external possession (cf. Russian: *On mne slomal ruku* 'He broke my arm'), beneficiary (cf. Russian: *On mne kupil knigu* 'He bought me a book'). The map below shows the boundaries of English *to* and the Russian dative in the area of dative functions.[13]

While Haspelmath's map zooms in on the particular functions in the dative domain,[14] some more general (external) links should be mentioned as well. Thus,

[13] Again, this map may turn out not to be exhaustive if it is complemented with cross-linguistic data. Creissels (p.c.) has made us aware of a probable connection between experiencer and beneficiary.

[14] This is of course not the deepest possible level of decomposition either; thus, individual functions such as experiencer turn out to be heterogeneous, as the research on the typology of verb types has revealed (Tsunoda 1981b; Malchukov 2005). Also some other links, not represented on the map, might be posited, for example, a direct link between beneficiary and purposive functions, as attested in many Australian languages that do not use dative case for R arguments (cf. Luraghi 2003: 17).

978–0–19–920647–6 35-Spencer-ch34 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 529 of 534 May 16, 2008 18:59

as is well known from languages displaying Differential Object Marking (e.g. many Indo-Aryan; cf. Kittilä and Malchukov, ch. 36), recipients and patients may be marked by the same ('dative–accusative') case. On the other hand, in many languages where dative aligns with directional, the polysemy can extend further into the locative domain (which is true for the French *à*, but also for many Altaic languages). Further, some languages (in particular, Australian and Austronesian) show dative–genitive polysemy, thus extending dative marking further beyond the beneficiary and external possession. Some of these connections will be pursued below.

34.6 Beyond core arguments: instrumentals and comitatives

A relatively large number of semantic maps have been proposed in the area of non-core participant roles prominently featuring comitatives and instrumentals, including Heine et al. (1991: 166), Michaelis and Rosalie (2000), Luraghi (2001), Haspelmath (2003), Yamaguchi (2004), and Narrog and Ito (2007; cf. also this volume, Chapter 40). An overview and comparison is provided in Narrog and Ito (2007). Stolz (1996b, 2001a, b, and others) has done the most extensive work on polysemy in this area, however, mostly without explicitly referring to semantic maps. The relationship between instrumental and comitative is cross-linguistically well established in this area, as is the relationship of comitative to noun conjunction and to possession (e.g. Stassen 2003; Stolz 2001b), and the relationship of the instrumental to ergative and other agents (cf. the work of Palancar 2002 etc.; also Stolz 2001a), to cause/reason, to manner, and to material. Research results in this area are represented on the map in Chapter 40 and are integrated into the general map in section 34.8.

34.7 Other domains: spatial forms

There are a number of other semantic domains which cannot be addressed here for reasons of space but should at least be mentioned. The most complex domain including a variety of functions displaying complex polysemy patterns is arguably the spatial domain (Stolz 1992; Creissels 2006a).[15] Apart from polysemies among the

[15] An original approach to the study of locative markers (adpositions) has been developed by S. Levinson and his associates at MPI for Psycholinguistics in Nijmegen (see e.g. Levinson et al. 2003).

functions within the spatial domain,[16] there are numerous links beyond the spatial and non-spatial domains, moreover the former may be conceived as source domains with respect to the latter, as explicitly acknowledged in localist case grammar (Anderson, Chapter 8 *et passim*) but also by Croft (1991). This is also manifested in diachronic developments of spatial markers to non-spatial functions as recently summarized in Yamaguchi (2004). In particular, Yamaguchi describes paths of semantic extension of spatial adpositions into the agentive domain, where some developments are direct (from ablative to ergative/agentive), while others are indirect (thus extensions from locative and prolative to agentive are mediated by instrumental function, from allative to agentive are mediated by recipient, beneficiary, or purpose functions). Such developments are also familiar from grammaticalization literature (see section 34.3 above). Another complex domain, largely disregarded in this presentation, is the possessive domain,[17] which, in its turn, is connected to other domains, in particular, to the spatial one through different conceptual scenarios (called 'grammaticalization schemes' by Heine).

34.8 Towards a general map of case functions

One advantage of semantic maps is that partial maps represent parts of a general network, and thus can be integrated with each other (Haspelmath 2003). In this section (see Figure 34.8) we present a tentative general map for the major (non-spatial) functions, which is based on the previous literature, and at the same time summarizes research of polysemies involving individual cases in contributions to this volume.[18]

The upper part of the map linking agents to instruments and causes, on the one hand, and to genitives, on the other, finds support in the most common polysemy

This approach is similar in spirit to the semantic map approach but differs in methods of data collection, informed by psycholinguistic research.

[16] One generalization which emerges from the study is that there are topological constraints on polysemies involving goal–(static) location and source markers to the effect that many languages where these markers are polysemous conflate goal and location, some all three functions, still others source and goal, but apparently there is no language where the same marker encodes goal and source to the exclusion of location (cf. Creissels 2006a).

[17] In the possessive domain the major distinction is between the possessor (POS) function (frequently encoded through genitives) and the proprietive function encoded by a special case marker in many Australian languages. The latter function, which is directly linked to the comitative–instrumental domain, will be disregarded here.

[18] The experiencer function is left out here. Being connected to a large number of other functions, its addition would not contribute to the informativity of the map.

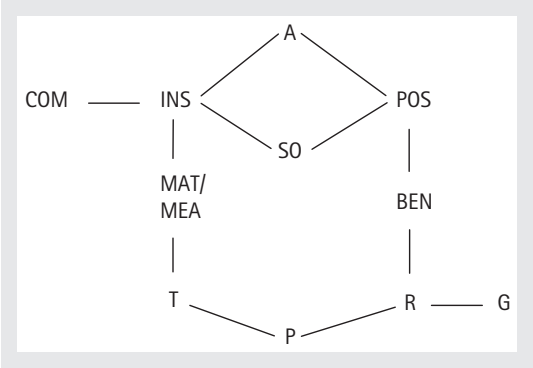


Figure 34.8. A general map for (major) thematic roles

patterns involving ergative case. A similar configuration has been proposed on semantic grounds by Grimm (2005), who noted that both 'routes' involve different reduction of the agentivity prototype: possessors and datives share the sentience property with agents but lack the feature of instigation, while instruments are instigating entities but lack sentience.[19] The lower part of the map is reminiscent of the ditransitive maps above, even though the T role is not conceived as restricted to ditransitive themes but also pertains to monotransitive themes (they differ from patients in that they do not necessarily involve a change of state in the course of a verbal event). Note that in contrast to alignment maps in section 34.2 the proposed map does not include a separate S function, since S does not correspond to a single macrorole. A concomitant difference is that the map presented here is concerned with the distribution of overt markers rather than zero markers, the distribution of which is governed by considerations of economy and local markedness (cf. Aristar 1997).[20] The upper and lower parts of the map are linked by two 'routes', one leading from Instrumental to Theme, and another from possessive to dative marking. The former route is mediated by the Means/Material function (cf. *load the cart with bricks*), which shows similarities to Instrument, on the one hand (may be conceived as an instigating entity), and with the Theme, on the other (standing in an incremental relation to the verb). This accounts for the fact that the Means/Material function may pattern either with Instruments or with Themes across languages, or indeed in the same language, resulting in a well-known 'spray/load alternation'

[19] Grimm (2005) used a modified list of Dowty's (1991) proto-agent properties (including instigation, motion, sentience, and volition) to construct an 'agentivity lattice', which models the gradual reduction of agentivity properties as one moves down the lattice.

[20] Indeed, zero marking is not confined cross-linguistically to intransitive subjects, but is also found in other cases of 'natural correlations' between a case function and the semantic content of a nominal. For example, in many languages, place names unlike other NP types may remain unmarked when used in a locative function, since this function is recoverable from the inherent properties of the place names (see Arkadiev, Chapter 47; Malchukov and de Swart, Chapter 22).

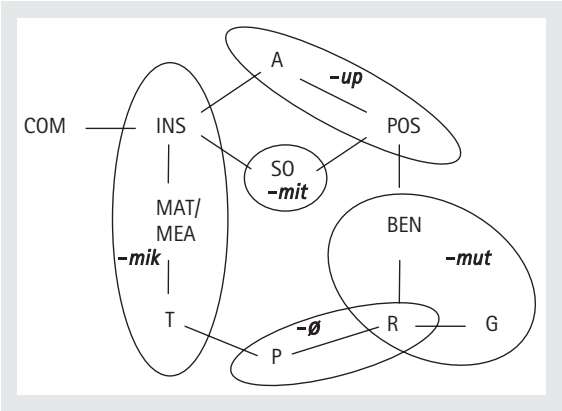


Figure 34.9. Case system of Eskimo (West Greenlandic)

(cf. *load bricks on the cart* vs. *load the cart with bricks*; see, e.g., Anderson 1977; Van Valin 2005 and also Bickel and Nichols, Chapter 20, for discussion). Finally, one of the conceptual underpinnings of the possessive–recipient connection may be the 'possession as goal' scheme, as suggested by Heine.

Thus the proposed map has both a certain semantic plausibility, but also finds empirical support in the frequent polysemy patterns, involving individual cases, as described in the contributions to this volume: the dative–allative polysemy, as familiar from English (recall the functions of *to*), is common across languages (see Creissels, Chapter 42; Næss, Chapter 38)[21]; the dative–genitive polysemy is attested in many Australian and Austronesian languages but is also found elsewhere (see Lander, Chapter 39; Næss, Chapter 38); the genitive–ablative polysemy is especially common in languages using adpositions for these functions (cf. Heine, Chapter 29, Lander, Chapter 39); the dative–accusative polysemy is familiar from languages with differential object marking (Næss, Chapter 38; Malchukov and de Swart, Chapter 22); instrumental–accusative polysemy is typical for languages with secundative alignment (see Kittilä and Malchukov, Chapter 36); the instrumental–comitative polysemy is the most frequent polysemy pattern involving both cases (Narrog, Chapter 40, Stolz, Stroh, and Urdze, Chapter 41); finally, the ergative–instrumental and ergative–genitive polysemies are identified as the two most frequent polysemy patterns involving ergative case (Palancar, Chapter 37).

By way of exemplification, consider the maps in Figures 34.9–10 showing a simplified picture of the functional range of (selected) case markers in Eskimo (West Greenlandic; Fortescue 1984), which is an ergative language with secundative

[21] In the cognitive literature this polysemy is conceived as instantiation of the 'Conduit Metaphor' (see Luraghi 2003).

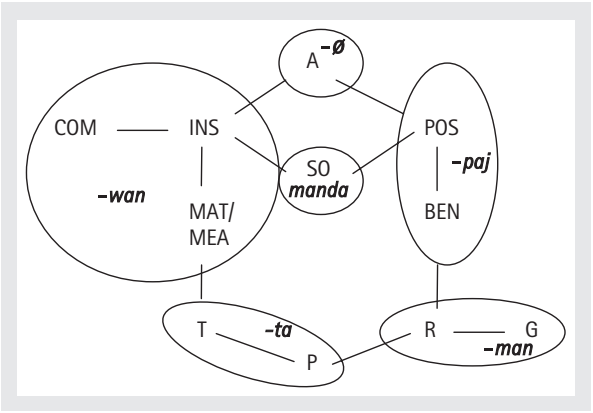


Figure 34.10. Case system of (Imbabura) Quechua

alignment, and Quechua (Imbabura Quechua, Cole 1982), which is an accusative language with indirective alignment.

It is clear that the general map in Figure 34.8 is in many respects incomplete. Thus it does not attempt a fine-grained functional analysis as exemplified by dative and instrumental maps. Instead it strives to outline a general picture. Some of the functions mentioned on the partial maps may be located as intermediate on the general map: thus, external possession can be arguably placed between adnominal Possession and Beneficiary, and Partitive function is intermediate between Source, Possession, and Theme. In some cases these missing intermediate functions can account for apparent discontinuities on the general map: thus, extensions of the genitive to the Theme–Patient domain may be mediated by the Partitive function. Further, the proposed map largely ignores the locative domain (e.g. the accusative–prolative syncretism is not represented on the map of case functions in Quechua), in particular, the (stative) locative function which shows multiple links across the map. This is also important, given that connections within the 'spatial dimension', which are not represented here in full, may also account for discontinuities on the map (e.g. in a language with a general locative marker used for Source, LOC, and Goal functions;[22] note that Source and Goal are not connected on our map).

Establishment of a more comprehensive and fine-grained map is a matter of future research. Yet already in the present form the map captures the major polysemy patterns for core and related functions, as well as making certain predictions about spread of meaning. As readers can verify, it provides routes for many diachronic developments noted in the grammaticalization literature (see section 34.3 above),

[22] This is attested in a number of Austronesian languages which have a general oblique (locative) marker with a wide range of functions. Thus in Manam, the general preposition *lo* has the following range of functions: Location, Goal, Beneficiary, Instrument, Source. This polysemy, which is difficult to account for in non-spatial terms (what is common to Instrument and Beneficiary?), is better analysed as different meaning extensions of the original polysemous locative marker.

spelling out where the spread is direct or mediated through other functions. It can also explain alignment patterns problematic for the alignment maps, like the double oblique pattern mentioned in footnote 5 (see also Chapter 47 by Arkadiev, and Chapter 48 by Stilo), which originates as an extension of the genitive–dative case to A and P functions.

34.9 Conclusions

In this chapter we have discussed one influential approach to the description and explanation of polysemy patterns across languages, the semantic map approach. We have further shown how different lines of research on case within a functional–typological tradition (alignment research, grammaticalization literature, cognitive approaches), can be unified within a semantic map approach. We also proposed a tentative semantic map of case functions focusing on non-spatial functions, which is designed to capture cross-linguistically recurrent case polysemy patterns.

In conclusion a note of caution is in order. Since semantic maps are functionally motivated they may be violated through interference of other functional, structural, or diachronic factors, and therefore should not be treated as absolute universals. Thus, violations of the semantic maps may either be due to historical processes of 'gram' replacement (see Haspelmath 2003 for examples)[23], or to phonologically conditioned syncretism, which need not form a contiguous segment of the map (e.g. dative–instrumental syncretism in Latin mentioned in section 34.1). It should be noted, though, that such violations are not necessarily fatal for semantic maps, since they are cross-linguistically rare and would be featured out on a larger sample.

Acknowledgement

Andrej Malchukov gratefully acknowledges the financial support of the Netherlands Organisation for Scientific Research (NWO) for the research reported here (grants no. 220-70-003 'Case cross-linguistically').

[23] See Narrog and Ito (2007) for a discussion of the instrumental marker in Kayah Li. Note, however, that in many, perhaps most, cases, case evolution and case displacement is partial insofar as the previous meaning is retained, thus causing no contiguity violations.

CHAPTER 35

MARKED NOMINATIVES

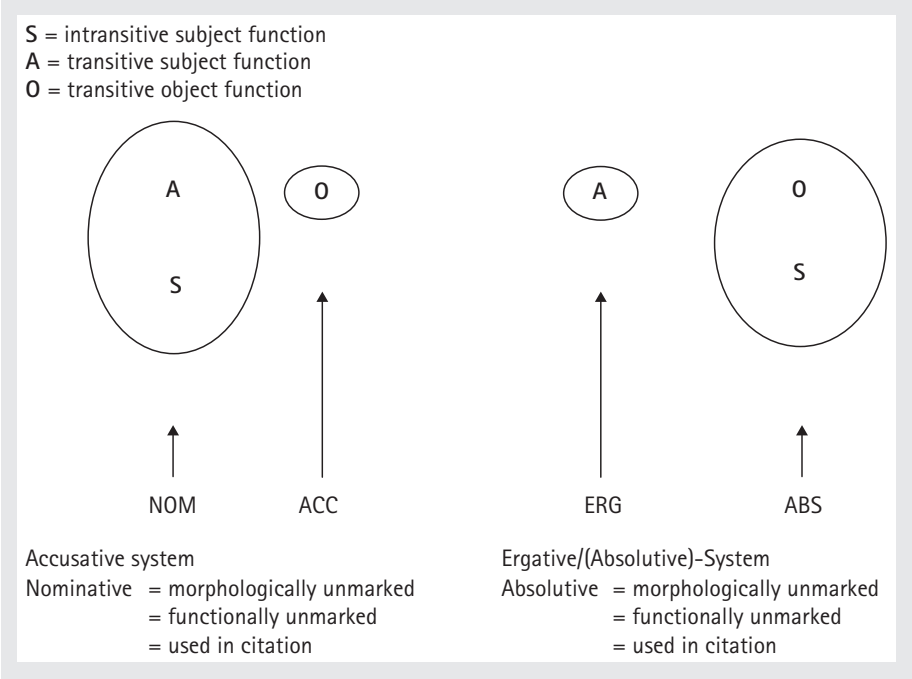
CHRISTA KÖNIG

35.1 Introduction

The following is mainly based on König (2006, 2008a, b). Due to space limitations I will not be able to give all details for the claims made here, and the reader is referred to those works.

As pointed out by Dixon (1994: 64f.), marked-nominative languages are a mixture of ergative/absolutive and nominative/accusative-systems (accusative in short). The pattern of the transitive subject, A, the intransitive subject, S, and the transitive object, O, is the same as in accusative languages, namely A and S are treated the same and simultaneously differently than O. They share this feature with accusative languages. I will use the term nominative for cases covering A and S, and accusative for cases covering O. But the accusative in marked-nominative languages is the morphologically unmarked form, at least typically (see below); it is used in citation, and it is also functionally the unmarked form. The nominative on the other hand is the morphologically marked form in a marked-nominative system; A, the transitive subject, therefore is encoded by the morphologically marked form. This feature of marked-nominative languages is shared with ergative systems.

With 'morphologically unmarked' I mean zero realization, while 'morphologically marked' means morphologically non-zero realization. 'Functionally unmarked' means being used in a wide range of different functions, often being the default form. 'Functionally marked' means being used in few functions only, not



Au: Please provide citation for figure 35.1.

Figure 35.1. Case systems

being the default form. The morphologically unmarked form is sometimes called 'basic form' in the literature. The morphologically marked form is derived from the morphologically unmarked form by adding some extra element. The morphologically unmarked form is shorter and/or underived vis-à-vis the morphologically marked form.

Theoretically one could also argue that accusative and ergative systems are a mixture as they share the feature that S is morphologically unmarked. However this feature is the expected one, as can be seen in Greenberg's universal 36: 'Where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb' (Greenberg 1966: 95). The oddity of marked-nominative systems can be seen in the fact that they violate Greenberg's universal 36.

Within accusative languages, the nominative is typically encoded in a morphologically unmarked form but there are some languages where both case forms, nominative and accusative, are morphologically marked. Two subtypes therefore need to be distinguished among the accusative languages.

In a similar fashion, with regard to the morphological markedness of nominative and accusative, two subtypes among the marked-nominative languages are to be distinguished: type 1 (the more common one), in which the accusative is the morphologically unmarked form and the nominative the morphologically marked form; and type 2, in which both case forms, nominative and accusative, are

morphologically marked. In type 1 of marked-nominative languages, the accusative is morphologically unmarked, functionally unmarked and used in citation. In type 2, the accusative is morphologically marked, but functionally unmarked, and used in citation.

Of the three properties used above to define a particular case system, namely morphological markedness, functionally markedness, and citation form, one is criterial, namely functional markedness. In a prototypical case system all three properties apply. The degree to which a marked-nominative system differs from an accusative and an ergative system varies. Type 1 marked-nominative languages share more properties with ergative languages than type 2 marked-nominative languages.

In sum, marked-nominative languages are defined as follows: A marked-nominative language is present when at least two cases are distinguished, namely an accusative covering O, and a nominative covering S and A. The accusative must be the functionally unmarked form; it is the default case, that is, the case which is used with the widest range of functions. If one of the two cases is derived from the other, it must be the nominative which is derived from the accusative and never the other way round. Type 1 shows more marked nominative properties than type 2.

In König (2008a) I have classified a language as a marked-nominative language even if it is only defectively so, meaning that there are neutralizations where there is no case distinction. Conditions for neutralization are, for example, definiteness (Wolaitta of the Afroasiatic phylum), person (Datooga of Nilo-Saharan; see Kießling 2001), noun phrase structure (Dinka of Nilo-Saharan; see Andersen 2002), gender (Cushitic languages of Afroasiatic), number (Cushitic languages), or constituent order (Suri-Chai of Nilo-Saharan; see Last and Lucassen 1998). If a marked-nominative language shows such a defective system I refer to it as a split language.

In East Africa a somewhat confusing terminology has been used: the morphologically unmarked form is called absolute or even absolutive (see e.g. Woldemariam 2003: 64, Tucker and Bryan 1966: 14, Leyew 2003: 237ff., Dimmendaal 1983, Last and Lucassen 1998, Hayward 1984), as it is called in ergative languages. I will follow Blake (1994: 26) who restricts the term absolutive to grammatical relations subsuming S + O (= P in his terminology) and nominative to all other options, including S + A. In order to be consistent I will consequently use the term nominative for a case covering A and S, irrespective of whether it is the morphologically marked or unmarked form (for further discussion see König, 2008a).

35.2 Case studies

I will illustrate the way marked-nominative systems work by a few case studies illustrating different types.

978–0–19–920647–6 36-Spencer-ch35 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 538 of 548 May 16, 2008 18:59

Table 35.1. Nominative and accusative case functions in Tennet

Case	Function
NOM	Subject (S & A) after the verb (see 1, 2)
ACC	(a) Citation form (6) (b) O (1) (c) Nominal predication (3) (d) Subject (S & A) before the verb (sometimes) (7) (e) Possessor in juxtaposed possession (5) (f) Peripheral participants introduced by head-marking devices (verbal derivation) such as the applicative (4)

35.2.1 Tennet

Tennet, a Surmic language (Nilo-Saharan), is an example of a canonical marked nominative language. According to Randal (1998) the accusative is the morphologically unmarked form. All other cases are marked by suffixes or tone. The nominative case is the morphologically marked form, expressed either by the suffix *-ɪ*, *-a*, or by tone.

A profile of the functions covered by the nominative and the accusative is presented in Table 35.1. Examples are referred to in parentheses. The nominative covers S and A, the accusative covers O, peripheral participants introduced by verbal derivation such as the applicative, nominal predicate in copula clauses, and possessor in juxtaposed possession, and is used as the citation form. In other words, one could say the accusative is the default case to encode dependent nouns if no special case is required. The accusative has a much wider range of uses; therefore it is also functionally unmarked.

(1) ákát *Lowór-ɪ* *Yomá.*
PFV.spear Lowor-NOM Yoma.ACC
'Lowor speared Yoma.' (Randal 1998: 231)

(2) *ók* *mányúdí-ɪ* *mgínaatí.*
go.PFV squirrel-NOM there
'Squirrel went there.' (Randal 1998: 231)

(3) *k-eénj* *anná* *deméz-zóh-t.*
1-be 1SG.NOM teach-AGEN.ACC-SG
'I'm a teacher.' (Randal 1998: 233; 2000: 72)

(4) *i-ttón-êk* *Lokóri-ɪ* *Yomá* *kavɪyâk.*
PFV-send-APL Lokor-NOM Yoma.ACC news.ACC
'Lokori sent news to Yoma.' (Randal 1998: 244)

(5) *k-ɛ-tɛ́ɛ́d-a* *ulúg-t* *ɔɔ.*
1-PFV-cut-1SG fish-SG.ACC head.ACC
'I cut the head of the fish.' (Randal 1998: 241)

Table 35.2. Nominative and accusative case functions in Maa

Case	Function
NOM	(a) Subject (S & A) after the verb (8 and 9) (b) Agent of passive[a] (18b) (c) Vocative (17) (d) Peripheral participants introduced by the preposition *tɛ̀*
ACC	(a) Citation form (10) (b) O (7) (c) Nominal predicate (see 'Sironka' in 11) (d) Subject (S & A) before the verb (16) (e) Possessor (13) (f) Peripheral participants introduced by head-marking devices (verbal derivation) (see 'father' in 14) (g) Patient (S) of passive[b] (see 18a) (h) After preposition *ɔ́* (12)

[a] The passive construction in Maa shows some irregularities due to the source construction from which it is historically derived. The passive marker is the third person plural pronoun 'they'; see Greenberg (1959), Heine and Claudi 1986: 79–84), König (2008a).
[b] see preceding footnote.

(6) *kʌvɪyâk.*
news.ACC
'News.'

(7) *Lɔkʊ́lɪ cí á-rʊ́h Lɔhâm.*
Lokuli.ACC AM IMPV-beat Loham.ACC
'It is Lokuli who is beating Loham.' (Randal 1998: 261).

35.2.2 Maa

The Nilotic language Maa (Nilo-Saharan) illustrates a marked-nominative language with case inflection being expressed exclusively by tone. The accusative is the basic form, and the nominative can be derived from it by complex tone rules (see Tucker and Mpaayei 1955). A profile of the functions covered by the nominative and the accusative is given in Table 35.2. Examples are given in parentheses. As can be seen, the accusative is the case with the wider range of functions. This is so even if the nominative encodes more than just S and A. Therefore in Maa as well, the accusative is the case which is not only morphologically but also functionally the unmarked one.

Maa (East Nilotic, Nilo-Saharan)

(8) *ɛ́-dɔ́l-íta ɔl-kítɛ̀ŋ en-kóítói.* V A O
3.SG-see-PROG M.SG-OX.NOM F.SG-road.ACC
'The ox sees the road.' (Tucker and Mpaayei 1955)

978–0–19–920647–6 36-Spencer-ch35 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 540 of 548 May 16, 2008 18:59

(9) *e-kuet-ita ɔl-kítɛ̀ŋ.* V S
3.SG-run-PROG M.SG-OX.NOM
'The ox is running.' (Tucker and Mpaayei 1955: 7)

(10) *ɔl-tʊŋánì*
M.SG-person.ACC
'person.' (Tucker and Mpaayei 1955: 175)

(11) *á-rá Sirónkà.*
1.SG-COP Sironka.ACC
'I am Sironka.' (Tucker and Mpaayei 1955: 175)

(12) *á-dɔ́l ɪn-kɛ́rá ó n-kitúààk*
1.SG-see F.PL-child.ACC PREP F.PL-woman.ACC
'I see the children and the women.' (Tucker and Mpaayei 1955: 215)

(13) *é-ípot ɔl-cɔ́rɛ l-ɔ́ [o]l-ayíónì*
3.SG-call M.SG-friend.NOM M.PEE-M.SG.POR M.SG-boy.ACC
'The friend of the boy called him.' (Tucker and Mpaayei 1955: 213)

(14) *á-ból-ókì papá ól-bɛ́nɛ.*
1.SG-open-APL father.ACC M.SG-basket.ACC
'I open the basket for father.' (Tucker and Mpaayei 1955: 129)

(15) *á-dúŋ-íé ɛnk-álɛ́m.*
1.SG-cut-INS F.SG-knife.ACC
'I cut it with a knife.' (Tucker and Mpaayei 1955: 142)

(16) *en-tító na-dɔ́l nɪnyɛ́.*
F.SG-girl.ACC REL.F.SG-see 3.SG.ACC
'It is the girl who sees him.' (Tucker and Mpaayei 1955)

(17) *lɔ́ tʊŋani.*
VOC person.NOM
'O person!' (Tucker and Mpaayei 1955: 176)

(18) a. *é-ísís-í Sirónkà.*
3.PL.A-praise-PAS Sironka.ACC
'Sironka is praised.' (Tucker and Mpaayei 1955: 175)
(Lit.: 'They praise Sironka.')

b. *e-rik-i [i]nk-ishu aainei* V S PP (Agent)
3.PL.A-lead-PAS F.PL-COW.ACC 1.SG.F.PL.POSS
[ɪ]l-mʊrran.
M.PL-warrior.NOM
'My cows will be led by the young men.' (Tucker and Mpaayei 1955: 81)

Marked-nominative languages are frequently split systems. The most widespread split condition has been described under the slogan 'no case before the verb', meaning that before the verb there is no case distinction. All core cases appear preverbally

in one case form only which typically is the accusative for marked-nominative languages. This split holds for all verb-initial and verb-medial marked-nominative languages. For obvious reasons, it does not apply to verb-final languages; it is an areal split characteristic of East Africa. It is not restricted to marked-nominative languages, it occurs as well in other case languages of East Africa. In Maa this split is also present: S and A occur only after the verb in the nominative (see 8 and 9), before the verb they are encoded in the accusative (see 16). As has been argued in König (2006), this irregularity has been the result of pragmatic constructions, a topic or focus construction. In Maa, the clause-initial position is used for focus participants. In the focus construction, the relative clause is still visible in the presence of a relative clause marker (see 16).

I have argued in König (2006) that the East African case languages probably have required the 'no case before the verb' rule as a result of historical processes. Originally the languages considered have been verb-initial, as most of them still are today, and they were case languages. In focus constructions, the focused participant is presented in a bi-clausal cleft construction with a preceding copula clause, with the focused participant being the nominal predicate, and the subsequent relative clause expressing the main clause semantics. Due to semantic pressure, the bi-clausal construction has been grammaticalized to a mono-clausal construction, roughly as follows:

Source structure: It is X who does Y. (With X being focused.)
Target structure: X does Y.

The copula clause–relative clause structure has been reinterpreted as the new main clause structure with a verb-medial order, but the cases of the source structure have been retained. As in all case languages, both ergative or marked nominative (or accusative), typically the nominal predicate occurs in the morphologically unmarked form – that is, the accusative in marked nominative languages and the absolutive in ergative languages; the rule applies to all case systems.

35.2.3 Datooga[1] (South Nilotic, Nilotic, Nilo-Saharan)

The Nilotic language Datooga is amongst the marked nominative languages where the accusative covers the widest range of functions: Ten different functions are covered by the accusative as opposed to one being covered by the nominative (see Table 35.3).

Datooga is a marked-nominative language with a split system: It either shows a marked-nominative system or no distinction at all. In addition to the 'no case before the verb' split (see function d in Table 35.3), Datooga also shows a person

[1] The following is based on an unpublished manuscript of Roland Kießling (2001) to whom I am very thankful.

Table 35.3. Nominative and accusative case functions in Datooga

Case	Function
NOM	subject (S & A) after the verb
ACC	(a) citation form (b) O (c) nominal predicate (d) subject (S & A) before the verb (e) possessor (f) peripheral participants introduced by head-marking devices such as verbal derivation (g) IO (h) after prepositions (i) S and A, first and second person, all positions (j) nominal modifiers in NP

Table 35.4. Personal pronouns in Datooga (Kießling 2001)

Person	Accusative	Nominative
1SG	áníinì, ání	
2SG	áŋìiŋì, áŋíɪ	
3SG	nìŋ	níŋ
1PL	éesèesá, èesà	
2PL	óogòogá, òogà	
3PL	**sàawà**	**sáawá**

split: According to Kießling (2001), the split is triggered by person, in first and second person the case distinction is neutralized. All core participants occur in the accusative. Table 35.4 shows the case forms of the independent personal pronouns in Datooga. The ones listed in bold are case-sensitive. They follow the case inflection rules, which apply to nouns as well: case is encoded exclusively by tone. The nominative is derived from the accusative by changing the initial and final vowel to a high tone.

35.2.4 Haro (East Ometo, Omotic, Afroasiatic)

The Omotic language Haro (Afroasiatic) is a type 2 marked-nominative language with a definite split system: according to Woldemariam (2003) case is restricted to definite nouns only. Of the ten cases, the three core cases are encoded by suffixes: The nominative, marked by the suffix *-i*, the accusative, marked either by the suffix *-a* (masculine, plural), or *-o* (feminine, singular) (called absolutive by

Woldemariam 2003), and the genitive by the suffix *-i*. Type 2 means there is no morphologically unmarked case; all cases, including nominative and accusative, are morphologically marked. Functionally, the accusative covers the widest range of functions, such as O (see 19), and the encoding of nominal predicates in copula clauses (20). The morphologically unmarked form in Haro presents the indefinite form of the noun, which is 'caseless'. The 'caseless' form is used with indefinite nouns in all slots which require the accusative, nominative, or genitive. Definiteness and case are both marked by separate suffixes which follow the noun stem. Nominative, accusative, and genitive case forms are built on the following pattern:

Noun – DEF – CASE.

(19) Haro[2]

ʔís-í garmá-z-a *ʔí-wod-ín-e.*

she-NOM lion-M.DEF-M.ACC 3FS-kill-PAST-A.DECL

'She killed the lion.'

(20) *yéʔ-í* garma-z-á-kko

that-NOM lion-M.DEF-M.ACC-FOC

'That is the lion.' (Woldemariam 2003: 64)

Type 2 marked-nominative languages are areally and genetically motivated, they occur in Highland East Cushitic, and in some Ometo languages (Omotic), all spoken in South Ethiopia, usually in languages where the encoding of case, gender, and number is amalgamated.

35.3 General features

The number of cases distinguished in a marked-nominative language varies between two (e.g. Umbundu, Bantu, Niger–Congo) and ten (e.g. Maale, Omotic, Afroasiatic).

The range of functions covered by the accusative in marked-nominative systems varies. The maximum range of functions found in all languages is fifteen. These functions do not all occur in one language only. There is a minimal list of functions basically covered by all accusatives in marked-nominative languages, which are (a) citation form, (b) object function, and (c) nominal predicate in copula clauses. In addition, the accusative may cover functions like (d) encoding indirect objects (e.g. in Arbore [Cushitic], Datooga [Nilotic], Umbundu [Bantu, Niger–Congo];

[2] As mentioned above, Woldemariam uses the term absolutive instead of accusative. In order to be consistent, the original glosses have been changed here.

see Schadeberg 1986), (e) being the basis for case doubling[3], a function which occurs in type 2 languages only, all of which are Afroasiatic languages (e.g. the Ometo languages Maale, Zayse, Wolaitta, and the Cushitic language Alaaba), (f) encoding further peripheral participants (e.g. time in Alaaba [Cushitic], locative and agents of passive clauses in Umbundu), and the following functions in the Cushitic language Dhaasanac: (g) focused participants, (h) topicalized participants, (i) possessee, and (j) modified nouns, that is nouns in a noun phrase which are modified.

König (2006; 2008a) proposes some remarkable generalizations with regard to marked-nominative languages: Marked-nominative systems are extremely rare outside Africa; the only other languages where they are found are the Yuman languages of California (e.g. Maricopa [Gordon 1986], Diegueño, and Jamul Tiipy [Langdon 1970, Miller 2001]). Within Africa however, they are by far the most frequent: There are sixty-one marked-nominative languages, as opposed to twenty-eight accusative and four ergative ones, among them three with a split marked-nominative/ergative system (see map). The map (Figure 35.2) gives an overview of the African languages with a marked-nominative system. Only the most significant features are considered, such as the genetic classification, tonal case systems, constituent order, and the most frequent or significant split conditions such as ergativity, no case before the verb, and definite. Note that Ongamo is not listed as a split language due to lack of data.

Marked-nominative systems in Africa are genetically motivated. They appear in certain branches of the Afroasiatic, Nilo-Saharan, and Niger–Congo phyla: in Afroasiatic branches such as Berber (e.g. Tamazight, Kabyle, Shiilh, Tuareg; see Aikhenvald 1990, 1995), Cushitic (e.g. East Cushitic: Dhaasanac, Rendille, Somali, Bayso, Highland East Cushitic: Sidamo, Alaaba, Kambaata, K'abeena, Libido, Gedeo, Burji)[4], and Omotic (e.g. Ometo: Maale, Haro, Koorete, Zayse, Zargulla, Gamo, Kullo, Wolaitta, Yem, Bworo, Kefa)[5]; Nilo-Saharan, such as Nilotic (e.g. Päri, Jur-Luwo, Dinka of West Nilotic, Maa, Ongamo, Teso, Turkana, Toposa, Karimojong of East Nilotic, and Kalenjin, Datooga, and Omotik of South Nilotic)[6] and Surmic (e.g. Majang, Didinga, Murle, Tennet, Baale, Mursi, Chai)[7]; and Niger–Congo, such as some western Bantu languages spoken in Angola and adjacent

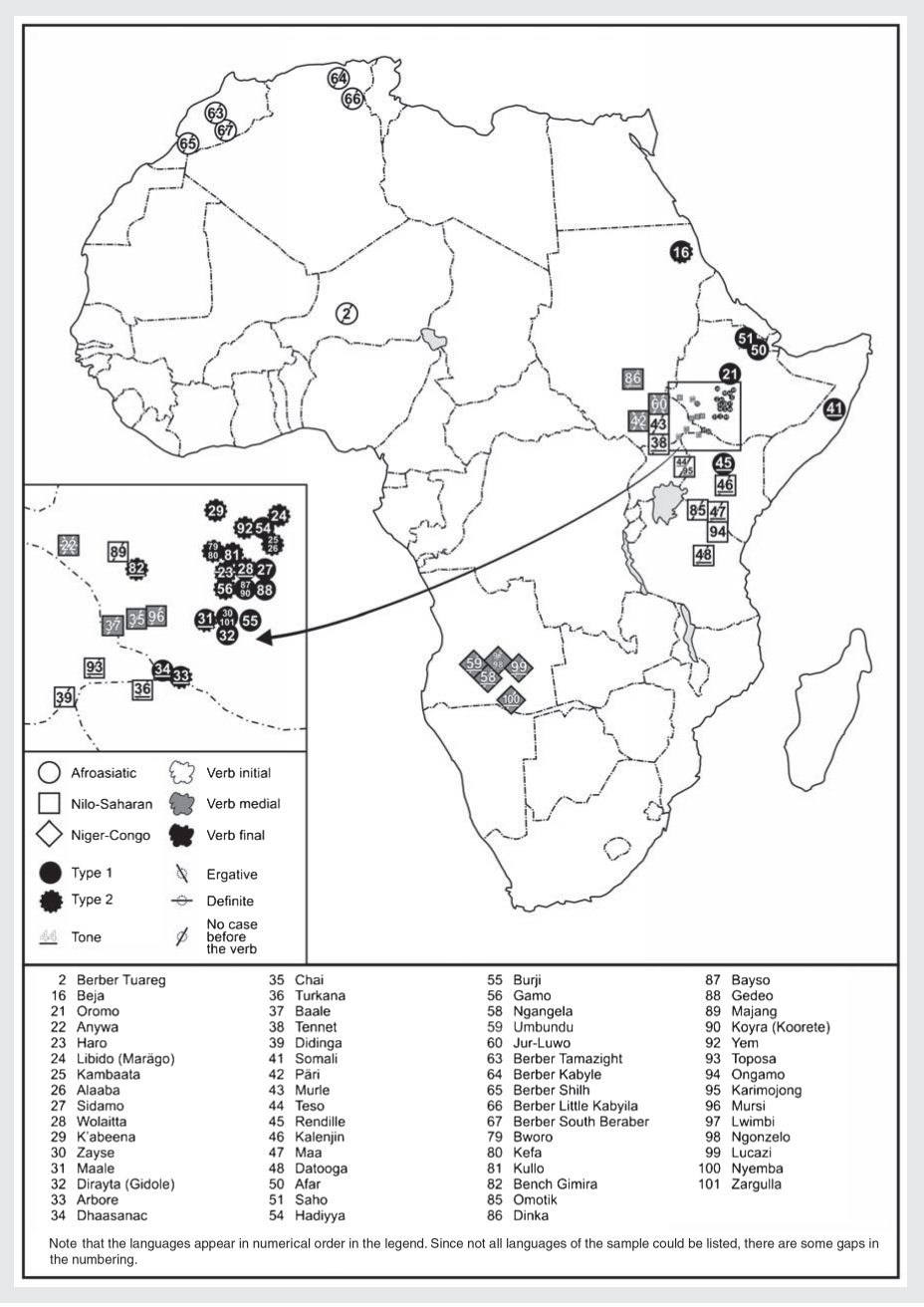
[3] In type 2 marked-nominative languages such as Maale, where locative cases are suffixed to the accusative form of the noun.

[4] For Dhaasanac, see Tosco (2001), for Bayso, see Hayward (1979), for Alaaba, see Schneider-Blum (2003), for K'abeena, see Crass (2005), for Burji, see Hayward (1988).

[5] For Ometo, see Fleming (1976), for Maale, see Amha (2001), for Haro, see Woldemariam (2003), for Zayse, see Hayward (1990).

[6] For Päri, see Andersen (1988, 2000), for Jur-Luwo, see Buth (1981), for Dinka, see Andersen (2002), for Ongamo, see Heine and Voßen (1975–6), for Turkana, see Dimmendaal (1983), for Toposa, see Schröder (2003), for Datooga, see Kießling (2001), and for Omotik, see Rottland (1982).

[7] For Surmic, see Dimmendaal and Last (1998), for Chai, see Last and Lucassen (1998), for Tennet, see Randal (1998).



2 Berber Tuareg
16 Beja
21 Oromo
22 Anywa
23 Haro
24 Libido (Marägo)
25 Kambaata
26 Alaaba
27 Sidamo
28 Wolaitta
29 K'abeena
30 Zayse
31 Maale
32 Dirayta (Gidole)
33 Arbore
34 Dhaasanac
35 Chai
36 Turkana
37 Baale
38 Tennet
39 Didinga
41 Somali
42 Päri
43 Murle
44 Teso
45 Rendille
46 Kalenjin
47 Maa
48 Datooga
50 Afar
51 Saho
54 Hadiyya
55 Burji
56 Gamo
58 Ngangela
59 Umbundu
60 Jur-Luwo
63 Berber Tamazight
64 Berber Kabyle
65 Berber Shilh
66 Berber Little Kabyila
67 Berber South Beraber
79 Bworo
80 Kefa
81 Kullo
82 Bench Gimira
85 Omotik
86 Dinka
87 Bayso
88 Gedeo
89 Majang
90 Koyra (Koorete)
92 Yem
93 Toposa
94 Ongamo
95 Karimojong
96 Mursi
97 Lwimbi
98 Ngonzelo
99 Lucazi
100 Nyemba
101 Zargulla

Note that the languages appear in numerical order in the legend. Since not all languages of the sample could be listed, there are some gaps in the numbering.

Figure 35.2. Marked-Nominative in Africa

areas (e.g. Ngangela, Umbundu, Lwimbi, Ngonzelo, Lucazi, Nyemba; see Blanchon 1998, 1999, Schadeberg 1986, and Maniacky 2002). Only a minority of marked-nominative languages are of type 2. In contrast to type 2 accusative languages, type 2 marked-nominative languages show a coherent picture. They appear just in one area which is located in southern Ethiopia, and genetically they belong to two branches of Afroasiatic only: they are either Ometo languages, in particular North Ometo, such as Gamo, Kullo, Wolaitta, Bench, Yem, Bworo, and Kefa, and South Ometo, such as Maale, Haro, and Zargulla; or they are Highland East Cushitic languages, such as Alaaba, Kambaata, K'abeena, and Libido. The coherent appearance of marked-nominative type 2 languages is probably due to the fact that case marking is intermingled with gender, or definiteness marking.

Marked-nominative systems are areally motivated, at least in East Africa. The western Bantu languages also form one area and all of them are genetically related. In the border region of Ethiopia, Sudan, Uganda, and Kenya there are neighbouring marked-nominative languages which are not genetically related, such as Dhaasanac and Turkana: Dhaasanac belongs to Afroasiatic, Turkana to Nilo-Saharan. They both encode the nominative by a low tone. This correspondence might be due to areal influence. Additional overlaps between the two phyla are present in the Surmic languages Majang, Murle, and Baale (all Surmic, Nilo-Saharan) and the Ometo language Bench (Afroasiatic), as well as in the Surmic language Chai (Nilo-Saharan; see Last and Lucassen 1998) and several Cushitic languages.

There is one additional feature with regard to marked-nominative which is typologically unique: in Africa there are languages where case is exclusively expressed by tone. As noted above, all tonal case languages are marked-nominative languages. In accusative languages, tonal case appears, if at all, only with pronouns, never with nouns.

So far, little is known about the origin of marked-nominative systems. According to Li et al. (1977), the Wappo system goes back to an earlier ergative system. Within Africa the following sources have been mentioned as the historical sources for the nominative case in marked-nominative languages: Topic (for East Cushitic, in particular Highland East Cushitic, and Ometo; see Tosco 1994: 234), a definiteness marker (for Anywa, Päri, Jur-Luwo [Northern Lwoo, Nilotic]; see Andersen 1988, Reh 1996; Berber [Afroasiatic]; see Aikhenvald 1990), and an agent encoding in passive-like clauses (e.g., for Maa, Dinka [Nilotic]; see Andersen 2002).

The following suggestions have been made on the origin of marked-nominative systems: Andersen (1988) claims that a marked-nominative system has arisen from ergative, although this claim has to be taken with care. For the Surmic language Tennet (Nilo-Saharan), Randal (2000) argues that the marked nominative is of ergative origin (but see König 2008a).

In the only split ergative/marked-nominative languages of Africa, Päri, Anywa, and Jur-Luwo (all West Nilotic), a development from a definite marker to a nominative case to an ergative case has been documented, suggesting a change from

marked-nominative to ergative system (see König 2008a). Hayward and Tsuge (1998) have argued that within Omotic languages (Afroasiatic), marked nominative is a recent development, arising out of an accusative system. Similar claims have been made by Tosco (1994) and Sasse (1984) for Cushitic. Taking the Californian language Wappo and the West Nilotic languages Päri, Jur-Luwo and Anywa into account, it is possible to argue that an ergative system develops into a marked-nominative system and that a marked-nominative system may develop into an ergative system, while evidence from Cushitic and Omotic languages suggests that a development from an accusative system to a marked-nominative system is also possible. In other words, marked-nominative can result overall from both ergative and accusative systems.

35.4 Conclusions

To conclude, marked-nominative systems represent a distinct type of case marking which is rarely found among the world's languages, while in eastern Africa it is extremely common, and in Africa it is by far the most prominent case pattern. It has been shown that all case systems expressed by tone in eastern Africa are marked-nominative throughout. Marked-nominative is not restricted to any constituent order: there are verb-initial languages, such as Turkana or Maa, verb-final languages, such as Dhaasanac or Somali, or verb-medial languages, such as Baale and Chai. The distribution of marked-nominative languages patterns to some extent with genetic boundaries, but there is also evidence that areal relationship, that is, language contact, must have played some role in its development. There is a high concentration of marked-nominative systems in the border region of Kenya, Uganda, Sudan, and Ethiopia, and the areal distribution of marked-nominative systems in this area cuts across the genetic boundaries that exist between Nilo-Saharan languages of the Nilotic and Surmic subgroups on the one hand, and Afroasiatic languages of the Omotic and Cushitic families on the other. There is some evidence to suggest that the direction of influence was from Afroasiatic to Nilo-Saharan languages; in certain cases, however, such as Dhaasanac and Turkana, it is plausible that at a later stage, areal influence also went from the Nilo-Saharan language Turkana to the Afroasiatic language Dhaasanac.

While marked-nominative systems represent a fairly consistent type among the world's languages, their development is complex. There are West Nilotic languages showing a split ergative or marked-nominative system. Historical evidence suggests that within Päri and Jur-Luwo, a development has taken place from an earlier marked-nominative to a later ergative case. Possible sources for the

978–0–19–920647–6 36-Spencer-ch35 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 548 of 548 May 16, 2008 18:59

marked-nominative are the expression of agents in passive clauses, and former topic markers. Berber languages also show a complex behaviour as they have not only marked-nominative systems but also split S as well. It seems that split S was already present in Proto-Berber. Marked-nominative is a later development, possibly arising out of an old definite particle.

The present chapter is a first attempt to give an overview of marked-nominative languages. In future research it would be helpful to compare the African type of marked-nominative languages with the few marked-nominative languages found elsewhere in the world. In particular, the following questions are of interest: First, is the profile of the accusative found in African marked-nominative languages similar to the profile of the accusative found elsewhere in marked-nominative languages? And second, are scenarios concerning the rise of marked-nominative languages applicable elsewhere in the world?

Further research should also focus on the pragmatic aspects of marked-nominative systems. Of special interest here is the question of whether topic functions are a possible source of marked-nominative case forms.

CHAPTER 36

VARIETIES OF ACCUSATIVE

SEPPO KITTILÄ
ANDREJ MALCHUKOV

36.1 Formal varieties of the accusative

36.1.1 Core meaning of the accusative

The core function of the accusative case is to encode the affected participant in a transitive clause. Examples of this are found below.

Finnish

(1) *henkilö* *tappo-i* *karhu-n* (*karhu-a)
person.NOM kill.PAST-3SG bear-ACC (*bear-PART)
'The person killed a bear'

(2) *henkilö* *ajattel-i* *karhu-a* (**karhu-n*)
person.NOM think-3SG.PAST bear-PART (*bear-ACC)
'The person was thinking about the bear'

(3) *henkilö* *jo-i* *maido-n/maito-a*
person.NOM drink-3SG.PAST milk-ACC/milk-PART
'The person drank the milk (ACC)/some milk (PART)'

Finnish is representative of nominative–accusative languages in that it shows that Patient objects of canonical transitive clauses as in (1) will be eligible for ACC

case-marking. Note that those objects that are not affected (as in (2)) or affected only partially (as in (3)), take partitive rather than accusative case in Finnish. In many languages ACC marking extends beyond semantically transitive clauses taking Agent and Patient arguments (such as 'kill' or 'break'), and therefore ACC is sometimes regarded as a purely structural case (especially in the generative tradition). Yet, a more balanced view, advocated by Blake (2001) among others, is to regard accusative as a syntactic case with encoding of Patients as its semantic core. Of course, affectedness, which is a defining property of a patient, is not the only feature that determines (ACC) marking of Os. We shall return to the role of other features such as animacy for object marking in section 36.1.3 below (see Kittilä, Chapter 23 and Malchukov and de Swart, Chapter 22, for further discussion of features contributing to semantic transitivity).

36.1.2 Formal varieties of the accusative

In (1)–(3) examples from Finnish were illustrated in which the accusative is expressed by attaching the suffix *-n* to the noun. This is the typical way of expressing the accusative in Finnish for nouns. However, the form of the accusative marker varies according to whether the element it attaches to is a noun or a pronoun. Thus in Finnish the accusative is formed with the suffix *-n* in the case of nouns (see (1)), while pronouns bear the suffix *-t* (e.g. *minu-t* 1SG-ACC 'me'). This variation is mandatory and the markers are not in free variation. The status of these affixes is also different, which means, for example, that *-n* is eliminated in passivization, while *-t* is not. These kinds of formal differences in the expression of the accusative (or direct object) are attested in many other languages, as Germanic languages such as English and Swedish, in which only pronouns have a distinct accusative form, illustrate. A somewhat similar case is illustrated in (4–5) from the Australian (Pama-Nyungan) language Bidjara:

Bidjara (Blake 1976: 282)

(4) *õura-õu munda bada-la*
dog-ERG snake bite-PAST
'A dog bit a snake'

(5) *õaya nuõu-na bada-la*
I.NOM he-ACC bite-PAST
'I bit him'

Bidjara is a language with split ergativity dependent on the nature of arguments: the marking pattern is absolutive–ergative for nouns and nominative–accusative for pronouns. This has the consequence that the direct object can be marked in two ways; it either bears an overt accusative marking (pronouns) or it bears no

overt marking (nouns). The variation is between an overtly marked accusative and a zero-marked 'accusative'. Generally, it is true that if a dedicated ACC case is found on nouns, it will be also found on pronouns, although exceptions to this do exist (Iggesen 2005b; see Iggesen, Chapter 16 for more discussion of 'case asymmetries' between different classes of nominals).

36.1.3 Restrictions on the use of the accusative

Given that pronominals frequently display idiosyncratic behaviour as compared to nouns, differential marking of pronominal and nominal Os discussed above may seem a merely morphological matter. Yet, as repeatedly observed in the typological literature (Silverstein 1976; Moravcsik 1978; Comrie 1981a; Bossong 1985a), this split is in fact a part of a more general cross-linguistic pattern, where overt accusative marking is restricted to animate (human) nouns, while inanimate nouns bear zero marking. This phenomenon known as 'Differential Object Marking' (Bossong 1985a) is illustrated below for Korku and Awa Pit:

Korku (Nagaraja 1999: 46)

(6) *iñj ga:Di(-ke) sege-pa*
I cart(-OBJ) bring-NONPAST
'I will bring a/the cart'

(7) *iñj siTa-ku-ke saya-ku-ba*
I dog-PL-OBJ take-PERS-NONPAST
'I will take the dogs'

Awa Pit (Curnow 1997: 72ff)

(8) *ishu=na pitikku ku-m*
tiger=TOP sloth eat-ADJZR
'Tigers eat sloths'

(9) *santos=ta=na mvza pyan-a-ma-t*
Santos=ACC=TOP almost hit-PL:SUBJ-COMP-PF.PTCP
'They almost beat up Santos'

In Korku, the affix *-ke* is mandatory with animate nouns, while it is only optional with inanimate nouns. In Awa Pit the ACC marker is possible only for human objects, as illustrated in (8)–(9). DOM displays a great deal of variation cross-linguistically in terms of what classes of nominals can (viz. should/must) be marked explicitly as (direct) objects, what are the features involved (animacy, person, definiteness), but availability of overt accusative marking is invariably restricted to the nouns occupying higher positions on the Animacy Hierarchy (see Malchukov and de Swart, Chapter 22 for discussion). In other languages ACC marking is sensitive to

978–0–19–920647–6 37-Spencer-ch36 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 552 of 561 May 16, 2008 18:59

relative prominence of subject and object on the animacy hierarchy. For example, in Yukaghir (Maslova 2003), ACC marking is absent if the subject is the first/second person, while it is obligatory if the subject is in the third person (i.e. when it is not higher than O on the person hierarchy). Ik is another language of this type (see König, Chapter 50).

Object marking may be further restricted if one looks (beyond monotransitive) at ditransitive constructions. Notably, some languages extend DOM to ditransitive clauses with animate objects (themes), while other languages do not. That is, in the languages of the former type animate Ts of ditransitives are marked in the same way as animate Ps of monotransitives, while in the languages of the latter type T is left unmarked even if animate. The former case is illustrated by Korku (in (10)), while the latter is illustrated by Awa Pit (in (11)):

Korku (Nagaraja 1999: 46)

(10) *ra:ja* *ra:ma-ke* *sita-ke* *ji-khe-nec*
king.NOM ram-DO Sita-DO give-PAST-PERS
'The king gave Sita to Ram'

Awa Pit (Curnow 1997: 72ff)

(11) *na=na* *santos=ta* *pashu* *mvla-ta-w*
1SG:NOM=TOP Santos=ACC daughter give-PAST-LOCUT:SBJ
'I gave my daughter to Santos'

In Korku, DOM is found both in monotransitive clauses (see (6)–(7) above) and in ditransitive clauses (in (10)). Note that marking of animate themes in the ditransitive clause (10) results in an identical marking of theme and recipient, a situation which is avoided in many languages for reasons of ambiguity. This also provides an explanation why in other languages (like Awa Pit in (11)) DOM is suspended in ditransitive clauses. Note that in Awa Pit, the O marker *=na*, which is obligatory on animate Os of monotransitives (see (9)), is not found on animate themes of ditransitives. (See Kittilä 2006a, b for further discussion of ambiguity resolution in ditransitives, and Malchukov 2008a for animacy effects in differential case marking, in general.)

36.1.4 Distribution of the accusative: accusative on non-typical hosts

The examples above where accusative markers attach to a direct object illustrate the canonical function of accusative case. In addition, accusative markers may also attach to other elements of clauses. For example, in many languages ACC

case is found on adverbials of distance and duration; cf. Finnish *juoks-i tunni-n/ kilometri-n* [run-3SG.PAST hour-ACC/kilometre-ACC] 'ran for an hour/a kilometre'. Even though accusative marks an adjunct rather than a direct complement, in terms of case-marking (only) the two constructions are similar. This similarity also carries over to accusative/partitive alternations (as in (3)), which is also attested for adverbials (see Maling, Chapter 5, for further discussion and exemplification). In some languages the accusative marker appears beyond time/duration adverbials. For example, in Arabic ACC is found on manner adverbs (cf. e.g. forms like *jiddiyy-an* 'seriously' (ACC) in Standard Arabic), and in Ge'ez it additionally marks nominal predicates (cf. *wa-kon=a nadāfe* [and-be.PRF=3SG archer-ACC] 'and he became an archer'; Weninger 1999: 39). A related pattern represents 'accusative of respect' as found in (Ancient) Greek; cf. examples like *diaphérein phúsin* 'to be different in nature(ACC)' (see Luraghi, Chapter 9).

ACC marking is common on nominalized complement clauses; cf. the following example from Even (Tungusic): *asi muchu-ri-va-n haram* [woman return-PART-ACC-3SG.POS know.I] 'I know that the woman returned' (Malchukov 1995). In some languages ACC appears on infinitives as well (e.g. in Quechua; see Spencer, Chapter 12, ex. 3). A more exotic example is provided by Koasati where variation between NOM and ACC on non-finite forms serves as means of switch-reference: the former is used as a same subject marker, the latter as a different subject marker (Kimball 1991: 226). Sometimes, deviant distribution of an ACC marker reflects its morphological status as a clitic rather than an affix. Thus in Samelai (Kruspe 1999: 262), an ACC marker is used as a proclitic on the postverbal object (which is reminiscent of the prefixal ACC marking), but is encliticized to the verb if the object is omitted.

36.1.5 Alternative strategies

In this chapter we shall use the term ACC (case) in a broad sense which will include, apart from case proper, also other forms of 'dependent-marking' including clitics (cf. enclitic *=ra* in Persian), particles (e.g. ACC postpositional particle *o* in Japanese), and adpositions (e.g. ACC postposition *=ko* in Hindi). The reason for grouping them together is that the distinctions between these categories and case proper are not clear-cut, and may depend on theoretical analysis. For example, the ACC marker *=ko* in Hindi can be analysed either as a postposition attaching to the oblique form of a noun, or as a case attaching to the oblique stem.

An alternative strategy for object marking is achieved through 'head-marking' (agreement/cross-referencing), which is opposed to different forms of dependent-marking (Haspelmath 2005a speaks of 'flagging' vs. 'indexing' of grammatical relations). Radically 'head-marking' languages, e.g. many Amerindian, lack case

978–0–19–920647–6 37-Spencer-ch36 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 554 of 561 May 16, 2008 18:59

distinction on (core) arguments altogether; here (object) agreement can indeed be conceived as an alternative strategy to (accusative) case. In other languages, which have both case and agreement, the two strategies may be used in combination, as exemplified for Amharic:

Amharic (Gasser 1983: 110)

(12) *girma bet-u-n gäzza-Ø(-w)*
PN house-DEF-ACC buy/PAST-3SG.I(-3SG.II)
'Girma bought the house'

(13) *girma bet gäzza-Ø*
PN house buy/PAST-3SG.I
'Girma bought a house'

Amharic is another language with a typical DOM system determined by definiteness distinctions: definite direct objects bear explicit accusative marking, while indefinite direct objects are zero-marked (see Amberber, Chapter 51). In addition to being overtly marked accusative, definite objects differ from indefinite ones in that only definite objects may be cross-referenced in the verb.

While verb agreement need not be complementary to case, word order is indeed an alternative strategy for differentiating between (Agent and Patient) arguments, which takes over when there are no formal differences between subjects and objects. Most illustrative in this respect is the phenomenon of 'word order freezing', that is, a change from a variable order of subjects and objects to invariable in cases when NOM and ACC happen to be non-distinct (cf. Neeleman and Weerman, Chapter 18). For example, in Finnish, nominative and accusative are morphologically distinct case forms in the singular (cf. *-ø* vs. *-n*), but this distinction is neutralized in the plural (both are marked by *-t*). The order of the arguments in these cases is invariably subject–object when both are in the plural, even though the order of arguments is free (pragmatically determined) elsewhere. Similarly in Russian, if the form of the subject and object happens to be identical (as is the case in certain declensional types), it cannot be changed without a concomitant change in interpretation. (cf. Jakobson's famous example *Matj ljubit dočj* 'Mother (NOM=ACC) loves the daughter (NOM=ACC)'). The relation between word order and case is also typologically confirmed by the correlation between the absence of case and the SVO word order (see Siewierska and Bakker, Chapter 19).

While agreement (viz. cross-referencing) and word order have been traditionally recognized as alternatives to case-marking of grammatical functions, other strategies should be mentioned as well. A number of languages (e.g. Chukchi) make liberal use of object incorporation (see Mithun 1984 for a general discussion). Tungusic languages illustrate the use of a lesser known 'reflexive' strategy of marking direct objects (Malchukov 1995; cf. Benzing 1955). Note that unlike other 'Altaic'

languages (e.g. Turkic), unmarked objects are generally not found in (North) Tungusic, unless the object takes a reflexive possessive marker. For example, in Even, direct objects usually take the accusative case in *-w/-u/-m*; however, in the presence of the reflexive-possessive marker *-j/-i/-mi* (in sg ∼ *-vur/-ur/-bur* in pl) the accusative marker is lacking (Malchukov 1995); cf.: *oro-m d'avran* [reindeer-ACC caught] 'he caught the reindeer' vs. *or-mi d'avran* [reindeer-REFL.POS caught] 'he caught his own reindeer'. The reflexive strategy makes sense functionally: since the noun in the reflexive-possessive form cannot be construed as a subject, reflexive marking helps to disambiguate subject from object (see Malchukov and Spencer, Chapter 45 for further discussion). A more exotic case of object marking is found in Gazi (Iranian), where pronominal markers ('agreement enclitics') consistently attach to the direct object and thus can be regarded as an unconventional ACC marking (Stilo, Chapter 48).

36.2 Functional varieties and polysemy patterns

36.2.1 Functional varieties of ACC

Some languages distinguish between definite and indefinite ACC cases. One option would be to use a Partitive case in the function of indefinite ACC, of the type we observed for Finnish in (3), or some other case expressing a partitive function, such as genitive in Russian: cf. *On vypil vodu/vody* [he drank water-ACC/GEN)] 'He drank the water/some water'. Evenki (Tungusic), has indefinite accusative in *-(j)A* (with allomorph *-o*) contrasted with the definite ACC in *-w(A)* (with allomorph *-mo*):

Evenki (Nedjalkov 1997: 193)

(14) *Oron-mo/Oron-o* *d'ava-kal*
reindeer-ACC/reindeer-ACCin catch-IMP.2SG
'Catch the/a reindeer'

The indefinite ACC can be also used in a partitive function (cf. *mokar-ve genne=* 'bring (some) firewood'; Nedjalkov 1997: 193), which highlights functional similarities between the two cases. Note, however, that while the Finnish partitive can be used to render (in)definiteness distinctions only in the case of plural objects (cf. *näin poiki-a* [saw boy.PL-PART] 'saw (some) boys'), the use of the indefinite accusative in Evenki shows no such restrictions. Another interesting use of the

indefinite accusative case is 'designative': if indefinite ACC appears in combination with a possessive marker, the object is understood as destined for a certain person (to which the possessive marker refers):

Evenki (Nedjalkov 1997: 147)

(15) *D'av-ja-v* *oo-kal*
boat-ACCin-1SG.POSS make-IMP.2SG
'Make a boat for me'

Designative cases in Tungusic are further discussed in Malchukov, Chapter 44.

As already mentioned in §36.1.3. above, in many languages (e.g. in most Uralic and Altaic) an object can be left unmarked if it is indefinite/non-specific, in a classical DOM pattern. Interpretation of these cases is controversial (see Iggesen, Chapter 16, and Baerman, Chapter 14, for somewhat different analyses of this pattern). Some grammars of Turkic and Mongolian languages qualify these forms as 'zero accusatives', while others qualify them as nominative. Yet another approach will distinguish between nominative forms with a zero exponent and nouns unmarked for case involved in a DOM pattern. The latter approach stands to reason given the similarity of constructions with unmarked Os to noun incorporation in other languages (see Malchukov and Spencer, Chapter 45, for further discussion). This is also consistent with the fact that 'zero accusatives' can appear in many languages only when adjacent to the verb. For example, in Turkic languages unmarked objects normally appear only in the preverbal position (see Johanson 2006 *et passim*).

36.2.2 Polysemy patterns[1]

36.2.2.1 *Polysemy involving core cases*

Expression of O in the same form as S is of course very common, as witnessed by ergative languages, where the absolutive form is used in both functions. Yet, in most cases the absolutive argument is unmarked, so it is less appropriate to speak of polysemy here. Yet in other languages, an (overt) ACC spreads to some intransitive predicates resulting in a kind of split intransitive (alias active/stative) system. Although split intransitivity is more frequently manifested through cross-referencing, in some languages it is manifested through case-markers as well. Consider the case of Quechua where experiencer subjects of some intransitives (desiderative and stative predicates) appear in the ACC:

[1] The polysemy of the accusative case markers has been recently discussed in Henkelmann (2006). This study came to our attention too late to be taken into account in the present chapter.

978–0–19–920647–6 37-Spencer-ch36 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 557 of 561 May 16, 2008 18:59

Quechua (Cole 1982: 107)

(16) *Juzi-ta punu-naya-n*
Jose-ACC sleep-DESD-3
'Jose wants to sleep/is sleepy'

In Quechua accusative experiencers qualify as subjects by the standard syntactic tests (switch-reference, raising, etc.), and therefore can count as 'non-canonical' (i.e. non-canonically marked) subjects (Hermon 2001). The origin of such 'extended accusative' systems is assumed to be either analogical extension motivated by functional similarity between P arguments and subjects of unaccusative intransitives (Plank 1985; Harris and Campbell 1995) or direct reanalysis of the transitive ('transimpersonal') pattern (Malchukov 2008a; cf. Mithun 2008; Donohue 2007).

Another case of polysemy involving ACC markers relates to the much discussed issue of the alignment strategies in ditransitive clauses as compared to monotransitive (Blansitt 1973; Dryer 1986; Siewierska 2004; Haspelmath 2004). As is well known, either theme (T) or recipient (R) of a ditransitive construction can pattern like the object-patient (P) of monotransitives. In the former case ('indirective' marking) the distinction is between direct (P = T) and indirect objects (R), while in the latter case ('secundative' marking) the distinction is between primary (P = R) and secondary (T) objects (Dryer 1986; see Chapter 33 for the terms 'indirective' and 'secundative'). Note that in the latter case, the same case (ACC = DAT) marks both Ps and Rs, provided that the language makes use of dependent marking. Generally, it has been noted (Siewierska 2004; Haspelmath 2004) that case (or 'flagging' in general) favours indirective alignment, while agreement ('indexing') favours secundative alignment. This is obvious from the existence of 'mixed' patterns where alignment is indirective in terms of case-marking but secundative in terms of (object) agreement (see Haspelmath 2005a on Bawm). For example, in Ostyak (Khanty), object agreement is invariably with the recipient (i.e. secundative), while case-marking may be either indirective (as in (17) where R is introduced by a preposition), or secundative (as in (18) where T appears in the oblique case combining locative and instrumental functions):

Ostyak (Nikolaeva 1999: 40)

(17) *Ma a:n juwan e:lti ma-s-e:m*
I cup John to give-PAST-SG/O+1SG
'I gave the cup to John'

(18) *Ma juwan a:n-na ma-s-e:m*
I John cup-OBL give-PAST-SG/O+1SG
'I gave the cup to John'

In the Vakh dialect of Khanty (Tereshkin 1961: 65), a special oblique case (*tvoriteljno-objektnyj*) is used for T, while P and R remain unmarked (if nominal). Generally, languages like Vakh Khanty which are consistently secundative in terms of case are rare. More common are situations when the same case that is used for R, is also found on some Ps, yielding a familiar DOM pattern. Differential object marking prevents a neat distinction between the indirective and secundative alignment patterns, yielding a kind of a 'split' (or 'dual'; Siewierska 2004) marking. A further complication, also noted above, relates to the fact that languages displaying DOM may either restrict it to Ps of monotransitive clauses or extend it to Ts of ditransitives as well. Polysemies involved in different patterns of ditransitive alignment are further discussed in Malchukov and Narrog, Chapter 34.

36.2.2.2 *Polysemy beyond core arguments*

Turning to other cases of polyfunctionality involving ACC markers, we may note that some languages use the same case for encoding of O and Possessors. Of course, this pattern is frequent when both Ps and possessor are zero-marked (e.g. in consistently head-marking languages), but is found also elsewhere. For example, ACC/GEN polysemy is widespread in Uto-Aztecan; cf. (Derdrick and Casad 1999: 129) from (Sonora) Yaqui: *María=ta kuúna* [Maria-ACC husband] 'Mary's husband'. GEN is frequently involved in patterns of differential object marking (Moravcsik 1978). On the one hand, it can be used in a partitive sense to mark indefinite/partitive Os; cf. Russian: *On vypil vody* [he drank water-GEN)] 'He drank some water'. Less usual is the opposite situation found in Finnish, where GEN=ACC marking of definite Os is opposed to the partitive (as in (3) above). Similarly the use of Instrumental is restricted as an O marker, either as a marker of T-arguments (as in some languages with secundative alignment), or as a marker of indefinite objects involved in a DOM pattern, as reported for Yukaghir and Eskimo.

36.2.2.3 *Polyfunctionality with locative cases*

We have already illustrated examples of accusative–dative polysemy, which is common in languages with DOM. This polysemy is attested in many language families (cf. Bossong 1985a), including Romance languages (cf. preposition *a* in Spanish), Indo-Aryan (the enclitic *=ko* in Hindi), Tibeto-Burman (cf. object suffix *=wa* in Kham), Afro-Asiatic (cf. preposition *lil* in Maltese Arabic), and in some South-American languages (e.g. *=ta* in Awa Pit in (9) and (11)). In many of these languages polyfunctionality extends beyond core cases when dative has additionally an allative function (cf. Næss, Chapter 38). Accusative–allative polysemy that is not mediated by the recipient function does exist (Creissels, Chapter 42), but in most cases seems to be lexically restricted (e.g. few motion verbs in Korean allow the Goal in the

accusative). Perhaps, more surprisingly, P can share the same marking with the Source as well. Thus in Iranian languages some ACC markers are arguably of ablative origin; cf. the accusative marker *=de* of the locative–ablative origin in Sangesari (Stilo, Chapter 48). Some other accusative markers of spatial origin have been reported for a number of languages: thus in Rumanian, the preposition *pe* 'on' is used to introduce animate objects, while in Quechua the accusative enclitic *=ta* also has a prolative function ('through X, past X').

Still other polysemy patterns are attested, but they seem to be more rare or restricted in different ways. Of course, in languages with an impoverished case system, polyfunctionality of the ACC (better characterized as an oblique) can extend to further functions, but then it is likely to include one or several of the functions mentioned above (see Arkadiev, Chapter 47).

36.2.2.4 *Diachronic dimension of polysemy patterns*

As is clear from the discussion above, in many cases polysemy patterns can be best viewed diachronically as gradual extensions from one function into another. Thus, allative–dative–accusative polysemy represents a well established grammaticalization chain (cf. Lehman 1995; Aristar 1996; Heine and Kuteva 2002: 103; see also Heine, Chapter 29). The origin of the accusative–genitive polysemy is functionally less transparent, except for the cases when GEN is used in a partitive function (cf. Lander, Chapter 39). In fact, in some instances (as in Finno-Ugrian) their identity is due to a phonologically conditioned syncretism. In some other cases the accusative–genitive polysemy seems to be mediated through the dative function. This holds for Uto-Aztecan languages, which have a secundative alignment in ditransitives, that is, use the same case for P and R arguments.

Another path of grammaticalization concerns ACC markers of verbal origin (cf. Blake 2001; Heine and Kuteva 2002: 289–90). Consider this example from Fongbe, where the object (theme) is introduced by the 'light' verb meaning 'take':

Fongbe (Lefebvre and Brousseau 2002: 445)

(19) *kòkú só àsón ó ná àsíbá*
PN take crab DEF give PN
'Koku gave the crab to Asiba'

Similar examples of the use of TAKE verbs in a serial verb construction to introduce the object are attested elsewhere; in particular, they are common in isolating languages of Africa and South East Asia (Lord 1982; Heine and Kuteva 2002: 289–90). As such a 'disposal' verb grammaticalizes it can eventually give rise to an object ('accusative') marker. A well-known example is (Mandarin) Chinese where the coverb/preposition *bǎ* historically derives from a verb meaning 'grasp' (Li and Thompson 1981; see also Enfield, Chapter 57, for discussion of a similar

'disposal' construction in Lao). Note that 'take' verbs in a disposal construction can independently develop into the instrumental marker (that is, a serial verb construction of the general form 'X take Y (and) (X) V-s Z' can be reanalysed as 'X with Y V-s Z') (Heine and Kuteva 2002: 290; Blake 2001). This can ultimately result in a secundative alignment with T being marked identically to Instruments.

Sometimes accusative case has its origin in pragmatic markers (cf. König, Chapter 35, on the pragmatic origin of 'marked nominatives'). Thus case prefixes in Berber arguably originate from prefixed determiners marking definite-topical NPs (Sasse 1984: 122; see Kulikov, Chapter 28). This grammaticalization path can be naturally explained under the assumption that direct objects are secondary topics (Givón 1984). On the other hand, we also find cases where ACC markers develop from reanalysis of the focus construction as suggested by Tosco (1994) for Cushitic (see also König, Chapter 35; Heine, Chapter 29). The case of Gazi (an Iranian language), where direct object is marked by 'agreement enclitics' seems to be a related pattern (see Stilo, Chapter 48, for discussion and exemplification). The possible motivation behind this path of reanalysis, which is in a way opposite to the one proposed by Givón, is that Os usually constitute a rhematic part of a transitive construction.

36.2.3 Use of case in complex (adpositional, etc.) structures

The range of functions of the ACC increases further if one takes into account its use in complex constructions. A pattern with an ACC used in adpositional phrases in a Goal function is familiar from Indo-European languages: here ACC contrasts with DAT in encoding of Goal as opposed to static location; cf. German and Russian: *geht in die Schule* [goes in the.F.ACC school.F.ACC] R. *idet v školu* [goes in school.ACC]. In some South-Slavic languages ACC has generalized its use giving rise to a General Oblique Case used in prepositional cases ousting other cases from this function (see Sobolev, Chapter 49). Altaic languages provide examples of the use of ACC with postpositions; usually this is due to the origin of the postpositional markers from (transitive) verbs; cf. Even: *d'uu-v ereli* [house-ACC around] 'around the house' (*ereli* ← *erel-* 'circle; turn round'). Similar cases are attested in Dravidian languages; cf. Tamil: *viiTT-ai viiTTu* [house-ACC from] 'from the house' (*viiTTu* ← *viTu* 'leave'). A somewhat similar pattern is found in languages where accusative form is identical to the oblique stem, and other cases are derived from it; cf. multiple case marking in Maale (Cushitic) where the emergent ablative locative case attaches to the 'accusative stem' (Amha 2001): *máár-ó-idda.ppa* [house-ACC-LOC-ABL] 'from the house'). This is parallel to a situation found in many ergative languages, where

978–0–19–920647–6 37-Spencer-ch36 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 561 of 561 May 16, 2008 18:59

the ergative case serves as a base for further case inflection (Palancar, Chapter 37; Daniel and Ganenkov, Chapter 46).

Acknowledgement

Andrej Malchukov gratefully acknowledges the financial support of the Netherlands Organisation for Scientific Research (NWO) for the research reported here (grant no. 220-70-003 'Case cross-linguistically').

CHAPTER 37

VARIETIES OF ERGATIVE

ENRIQUE L. PALANCAR

37.1 Ergative case[1]

Ergative case – also ergative marker or simply ergative – is the term given to the grammatical morpheme associated with the noun phrase (NP) functioning as subject of a transitive clause (i.e. the A syntactic relation in Dixon 1994).[2] In semantic terms, ergatives mark NPs that typically play the role of agents in the transitive event rendered by such clauses. Two examples of this core function of ergative markers are given in (1):

(1) a. Basque (LI) (Manandise 1988: 8; glosses adapted)
(*Miren=**e-k***) *liburu bat* *irakur-r-i* *d-u-Ø*
Mary=L-ERG book one.ABS read-L-PRF 3.ABS-*UKAN*('have')3SG.ERG
'Mary has read a book'

b. Tauya (P-TNG), (MacDonald 1990: 321)
(*ʔe* *fanu-**ni***) *fenaʔa-ra* *Ø-yau-a-ʔa*
DEM man-ERG woman-TOP 3SG-see-3SG-IND
'The man saw the woman'

[1] This Chapter was written under the auspices of CONACyT Project No 47475. I want to thank Yuri Lander for all his comments on an earlier version of this chapter. To Even Hovdhaugen, for his support and inspiration while I worked on this topic in Oslo. I can be contacted at Facultad de Lenguas y Letras, Universidad Autónoma de Querétaro, CU, Querétaro 76010, Mexico, or at epalancar@hotmail.com.

[2] The term 'subject' is used here in a loose way.

978–0–19–920647–6 38-Spencer-ch37 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 563 of 571 May 16, 2008 19:0

The morphemes *=ek* in (1a) and *-ni* in (1b) are instances of ergative case by virtue of being morphemes associated with the NPs working as subjects in the transitive clauses in (1), which also express the agents of the events described.

37.1.1 Ergatives that are not ergative case

Ergative case is one of two possible morphological mechanisms to encode an ergative–absolutive alignment. In such an alignment, A (transitive subject) is treated differently from S and O (intransitive subject and transitive object), which are commonly treated alike. As illustrated in (1), ergative case is used on the NP functioning as A; this is called a 'dependent-marking' strategy. The other way to encode such an alignment is using a so-called 'head-marking' strategy; that is, using a special pronominal morphology in the verb to cross-reference A, which is different from the one encoding S or O. Traditionally, this pronominal morphology has also been called 'ergative', but it should not be confused with 'ergative case'. There are ergative languages that only use a head-marking strategy to encode their ergative–absolutive alignment, for example Mayan languages, Mixean languages, and the Abkhaz-Abaza subgroup of North West Caucasian. An example of such a strategy appears in (2) in Mayan Awakatek (Mateo-Toledo et al., in press):

(2) *ja* *Ø-s-tz'am* (*Wa'n*)A (*tzee'*)O
COMPL 3SG.ABS-3**SG.ERG**-GRASP John tree
'John grasped the tree'

Notice that there is no ergative case on the subject NP *Wa'n* 'John'. Reference to the A relation is only encoded in the verb with *s-*. Because of this, constructions of the type in (2) will not be regarded here as instances of 'ergative case'.

37.1.2 Where is ergative case?

I base the observations in this chapter on the study of the behaviour of the ergative markers from about 140 languages. Much of the information comes from the analysis and understanding of the linguists who have described such languages. Ergative case is found in almost all ergative languages in Eurasia (N.B. the number of languages in parentheses represent the languages for which I had information about ergative case when compiling this chapter): in all Caucasian languages, except the Abkhaz-Abaza subgroup of North West Caucasian (18 languages); in the ergative languages of the Indo-Iranian subgroup of Indo-European (5); in all Tibeto-Burman languages (12); in a number of language isolates, such as Basque (1) and Sumerian (1); in a number of ergative Uralo-Yukaghir languages (1); and in Chukotko-Kamchatkan Chuckchee (1). In Oceania – including Papuan New Guinea, Australia, and the Pacific islands – it is found in many Papuan

978–0–19–920647–6 38-Spencer-ch37 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 564 of 571 May 16, 2008 19:0

languages (3); in all Australian languages that are ergative (16 Non-Pama Nyungan and 70 Pama-Nyungan; the sample comes mostly from Blake 1977); and in Austronesian (4 Oceanic). In the Americas, it is found in Eskimo (3), in a number of Zoquean languages from the Mixe-Zoque macrophylum (1), and recently, as descriptive work progresses in the area, it is being found in many South-American languages from various subgroups (Carib (2); Chibchan (1); Yanomam (1); and Panoan (1)). All in all, it appears that the number of ergative languages which employ ergative case as a morphological exponent of ergative–absolutive alignment far surpasses the number of languages which use head-marking patterns exclusively.

37.1.3 Formal varieties of ergative case

In phonological terms, ergative cases are in general highly grammaticalized morphemes: bisyllabic ergatives are scarce (e.g. both the allomorph *-njdja* of Djingili (AT-NPN) (Chadwick 1976) and Ngangikurrunggur *-niŋgi* (Hoddinott and Kofod 1976) should be considered rare typologically). Much more common are ergatives in /CV/ or /VC/, and even only /C/ (e.g. Kabardian (CA-NW) *-m*, Colarusso 1992); Basque (LI) = *k* (= *e-k*, after consonants); etc.). Interestingly, though perhaps irrelevant, ergatives in V also appear rarely (e.g. Khinalug (CA-E) *-i*,Kibrik 1994b).

Morphologically speaking, affixal ergative cases are more frequent than other morphological types, but this may just reflect the morphological typology of the language in question. When affixal, it is more common to find ergative cases of the agglutinative type than of the inflectional Indo-European-like type. In many languages of the Pacific, case is marked by adpositions, and thus ergative adpositions are found when the language has ergative–absolutive alignment (e.g. Tokelau (AU-Pol) *e te malo* [ERG SPC government] 'the government', Hovdhaugen et al. 1989: 49). Interestingly, nominal case in ergative languages is often encoded by means of phrasal clitics, and as clitics are syntactically-bound morphemes, this function fits the prototypical role of an ergative case. Consequently many ergative cases are clitics, and in this respect, ergative adpositions of the Polynesian type are syntactically equated with ergative clitics.

On the other hand, affixal or clitic-like, ergative allomorphy abounds, although it is commonly conditioned on morphophonological grounds. Nevertheless, lexical restrictions also occur. If the language allows ergative case with various types of nouns (animate vs. inanimate, etc.), the allomorphy is normally not conditioned by the semantic type of the noun; but in the Bzhedug and Shapsug dialects of Circassian Adyghe, proper nouns receive *-ə*, while common nouns receive *-m* (*Ahmed-ə* [Ahmed-ERG.PN] vs. *ps'as'e-m* [girl-ERG]; Zekokh 1969 in Kumachov et al. 1996). Similarly, the ergative case used with nominals is frequently the one

employed with pronouns and other determiners, but there is a tendency for ergative formations to become morphologically opaque with the latter. For example, Basque has transparent case in first and second person (*hi=k* [you.familiar=ERG] vs. *gi-zon=a=k* [man=DEF=ERG]), but in demonstratives, there is root suppletion (*hau* [this.ABS] vs. *hon=e-k* [this.ERG=L-ERG]). In other languages, the ergative marker is altogether different in demonstrative pronominal paradigms (e.g. Adyghe and Kabardian (CA-NW) both have *-m* for nominal ergative, but Adyghe has *-š'* for demonstratives (*a-š'* [he/that.M-ERG]), while Kabardian has *-bə* (*a-bə* [he/that.M-ERG]). In Eastern Caucasian languages, the ergative forms for first and second person have become suppletive or opaque, e.g. Lezgian *zun* [I.ABS] vs. *za* [I.ERG] (Haspelmath 1993a); or Tsova Tush *so* [I.ABS] vs. *as* [I.ERG] (Holisky and Gagua 1994).

37.1.4 Restrictions on ergative case

Since Silverstein's (1976) seminal article, it is well known in linguistic theory that the use of ergative case on a particular noun may be ruled by an animacy hierarchy. This is especially the case for Australian languages, where in many of them, personal pronouns, proper and kinship nouns follow a nominative–accusative alignment whereas ergative alignment is used for other nouns (animate or inanimate). Specific languages idiosyncratically set their own boundaries across this hierarchy. In most Caucasian languages, regardless of their affiliation, the first and second personal pronouns together with proper names more often than not lack an absolutive/ergative inflectional contrast. The same is true for plural pronouns (and nouns) in many ergative languages (e.g. Basque *hai-e-**k*** {yonder.PL-L-ERG/ABS.PL}).

37.1.5 Ergative case on non-typical hosts

Besides nouns, in a number of languages ergative case is also used with verbal forms to express adverbial meanings. This mechanism appears to be particularly productive in Caucasian languages, where ergative case is often labelled 'oblique'. An example of a purpose adverbial function is given in (3) for Ubykh (CA-NW) (Hewitt 2005: 130):

(3) (*'txə-məɕa* *Ø-'s-tɕ'a-w-tə-**n***) *'a-məɕa-ɕ^{w}a-ʁa*
write-read(.ABS) 3SG-1SG-study-FUT-Class.II-ERG/OBL DEF-school-to
s-k^{j}'a-n-Ø
1SG-go-PRS-FIN
'I go to school in order to learn to read and write'

978–0–19–920647–6 38-Spencer-ch37 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 566 of 571 May 16, 2008 19:0

37.2 Functional varieties of ergative case

Besides encoding the subject of a transitive clause, ergative case also serves as grammatical exponency of an agent. The semantic role of agent may be defined as the participant that instigates and bears control upon a given action. In the prototypical case it is construed as a human being, and is thought to have acted intentionally and volitionally (for a full discussion see Palancar 2002). Expressing an agent is perhaps the foremost characteristic of ergative case, and as such it is exploited in sundry ways across ergative languages.

In a number of Australian and Papuan languages, ergative case is optional in transitive constructions (Blake 1977, Foley 2000). Some of these languages use it for pragmatic purposes to disambiguate the agent from the patient in a number of circumstances. In Dani (P-W), the use of ergative case on an agent noun is associated with events contrary to expectation (Foley 2000: 375). See (4a) where the ergative is not used on the noun *ap* 'man' because normally men eat pythons and not the other way around, as it is the case in (4b). The use of ergative case in (4a) would emphasize the agentive role played by the man:

(4) a. (*ap*) *palu* *na-sikh-e*
man python eat-RM.PAST-3SG.SBJV
'The man ate the python'
b. *ap* (*palu-**nen***) *na-sikh-e*
man python-ERG eat-RM.PAST-3SG.SBJV
'The python ate the man'

In other languages where ergative case is grammatically obligatory in transitive constructions – besides possible aspectual splits that condition its application – the case may alternate with other encoding possibilities. The use of the ergative in such situations reinforces a reading of control, volition, intention, and/or responsibility. Citing various authors, Foley (2000) points out that the use of ergative case in Papuan languages often contributes to some subtle meaning where actors are depicted as wilful agents. Foley (2000: 375) mentions that 'in Folopa (Anderson and Wade 1988), certain verbs construed as agentive and willful, like "kill", always co-occur with subjects in ergative case, whereas non-agentive verbs, like "like", never do so. In between are verbs that can be construed either way, e.g. "do/say", "get" '. The same is true of Indo-Aryan languages (IE) like Urdu/Hindi with infinitival constructions (Butt and King 1991, Butt 2005: 15), as in (5):

(5) (*nadya=**ne**/=**ko***) *zu* *ja-na* *hɛ*
Nadya.F.SG=ERG/=DAT zoo.M.SG go-INF.M.SG be.3SG.PRES
'Nadya wants to go to the zoo' (Erg=Control / Dat=Experiencer)

978–0–19–920647–6 38-Spencer-ch37 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 567 of 571 May 16, 2008 19:0

Along these lines, it is very common that languages use ergative case on the NPs functioning as subject of intransitive (unergative) verbs to render the actors as volitional or having control. This encoding possibility gives rise to the emergence of active/patient systems (Mithun 1991), and it is found across a number of Tibetan, Indo-Aryan (Kachru 1987), and Eastern Caucasian languages.

37.2.1 Ergative case polysemy patterns

Ergative case is often used across languages to express other semantic roles. Nevertheless, in many languages ergative case expresses nothing else but the agent in a transitive construction. I call the latter cases 'asyncretic' and the former 'syncretic'. In principle, there is nothing typologically odd about asyncretic ergatives, as they are found in all families. Figures may give a useful indication: twenty-one out of a total of eighty-eight ergative markers in Australian languages are asyncretic (24 per cent). The same is true of Tibeto-Burman (four asyncretic vs. eight syncretic) and in Caucasian where the proportion is larger (seven asyncretic vs. twelve syncretic). At times we find the vestiges of old syncretisms that have long vanished. Basque asyncretic ergative *=k* used to encode cause with intransitive predicates. This old causal use still survives in sporadic, formal, idiomatic expressions (e.g. *hotz=a=**k** n-en-go-en* [cold=DEF=ERG 1SG-PAST-*EGON* 'be.located'-PST] 'I was cold') and in the lexicalized adjective *hotzakil* 'very cold' (lit. 'dead from cold' *hotz-a=k (h)il* [cold=DEF=ERG die.PRF]).

37.2.1.1 *Common syncretic patterns*

Among syncretic instances, instrument is by far the most common semantic category found in ergative syncretisms. Figures speak for themselves: all syncretic ergative markers in Australian also express instrument (save the Djingili (AT-NPN) allomorph *-njdja*, which is better seen as an oblique). The pattern is found in the vast majority of East Caucasian languages (Lak being an exception; N.B. North West Caucasian languages have a separate Instrumental marker, and South Caucasian ergatives are asyncretic). It is also found in Indo-Aryan Brokskat, Konkani, and Marathi, but not in Urdu/Hindi, which has its own Instrumental. It is widespread across Tibeto-Burman and in Papuan languages, and a number of South-American languages have it too, for example, Shipibo-Konibo and Sanumá (Borgman 1990).

Occasionally, when an ergative expresses instrument, it may also serve to express cause in intransitive clauses (e.g. 'they died from hunger'). The pattern is rather scattered, but appears in a number of East Caucasian, Tibeto-Burman, Papuan, and Indo-Aryan languages, as well as in isolated cases. It is illustrated in (6) from Sanumá (MJ) (Borgman 1990):

(6) a. Agent (*kamisamakö-**nö***) *hama sam_atöpö se kite*
we.EXCL-ERG/AGEN visitor 1PL.EXCL hit FUT
'We will hit the visitors' (p. 29)

b. Instrument (*kusiali a-**nö***) *sa ia pia kule*
spoon 3SG-ERG/INS 1SG eat intend PRS
'I am about to eat with a spoon' (p. 123)

c. Cause (*kamali te wasu-**nö***) *ipa ulu a*
high.fever 3SG deadly-ERG/CAUS my son 3SG
noma-so-ma
die-FOC-COMPL
'My son died from a deadly high fever' (p. 123)

Possessor is another semantic category that ergative markers also express cross-linguistically, but in overall frequency, the pattern is much less common than the syncretism with instrument. Ergative markers expressing a possessive relation are typical of Eskimo and Tibeto-Burman languages. It is illustrated in (7) in Ladakhi (TB) (Koshal 1979). The pattern is sporadically found elsewhere, as in Caucasian Lak (Van den Berg 2005) or in Mixe-Zoquean Chiapas Zoque (Faarlund p.c.). (N.B. A similar head-marked pattern is found in Mixean and Mayan languages.)

(7) a. Agent (*thug-gu-**yi***) *pəl-ldən-ni kə-ne şpe-čhẹ*
boy-ABS-ERG/AGEN Paldan-GEN from-ABL book-ABS
khyers
take.SIMPL-PRF
'The boy took the book from Paldan' (p. 75)

b. Possessor (*khyi-**yi***) *sŋə-mə-riŋ-mo duk*
dog-ERG/GEN tail-ABS-long-ABS be.PRS
'The dog's tail is long' (p. 74)

However, it is typologically rare to find an ergative case that expresses both instrument and possessor at the same time, only in some Tibeto-Burman languages is this pattern possible (e.g. Athpare *-ŋa* (Ebert 1997b), Limbu *-le* (Van Driem 1987), and Tibetan *-s* have it (Denwood 1999)). (N.B. Panoan Shipibo-Konibo has it too; Valenzuela 1997.)

37.2.1.2 *Uncommon ergative syncretisms*

Other syncretic patterns are less common or reveal syncretic phenomena that are yet to be uncovered in undescribed languages. In North West Caucasian languages with nominal case, the ergative case has been treated as an 'oblique' marker because of its divergent functions. Apart from their locative functions, which other syncretic ergative markers also have (see next section), in these languages ergatives are the case governed by adpositions; the case that marks the possessor in head-marked

978–0–19–920647–6 38-Spencer-ch37 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 569 of 571 May 16, 2008 19:0

possession relations; they cover adverbial functions with verbs (see 3); and they may even express a dative participant, as well as other demoted patients. The last two options are illustrated in (8) in Kabardian (CA-NW):

(8) a. (*l'ə-**m***) *tχəɬə-r* (*fəzə-**m***)
man-ERG/AGEN book-ABS woman-ERG/DAT
Ø-jə-ri-t-a-s'
3SG.ABS-3SG.DAT-3SG.ERG-give-PRF-ASSRT
'The man gave the book to the woman' (Kumachov et al. 1996: 100)

b. *ɕ'a:lʒ-r* (*tχəɬə-**m***) *Ø-j-aw-dʒa-Ø*
boy-ABS book-ERG/DEMOTED.PAT he-it-DYN-read-INTR.PRES
'The boy is reading the book' (Hewitt 2005: 124)

A somewhat similar phenomenon is found in Carib Tiriyó (Meira 1999), where the ergative marker is used to encode a dative participant as well as the causee in causative constructions.

37.2.1.3 *Local syncretisms*

Syncretic ergative markers often encode spatial categories. Ergatives in Australian languages may encode locative as well as instrument (seventeen out of a total of sixty-seven syncretic instances in sixty-five languages), but in this linguistic area, no marker appears to encode agent and locative alone, without instrument. The same pattern is found in Shipibo-Konibo (Pan), in Vach Khanty (UY), and in Manipuri Meithei (TB). All ergative (oblique) markers in North West Caucasian languages also express a locative relation. After locative, ergatives encoding spatial source (ablative) also occur, but more infrequently; they are especially typical of Tibeto-Burman and Papuan languages. As with the locative, all such markers encode instrument and cause as well (e.g. Dani (P-W) *-(n)en* (Bromley 1981); Tauya (P-TNG) *-ni* (MacDonald 1990); Athpare (TB) *-ŋa* (Ebert 1997b); Thakali (TB) *-se* (Georg 1996)). Other spatial cases are rare. Perlative ('through/by the park') is found in Indo-Iranian Konkani (Maffei 1986) and Marathi (Kashli Wali p.c.), but outside Indo-Iranian this syncretism is not found, as with the allative, which appears circumscribed to South-American languages.

37.2.1.4 *Non-existent syncretic patterns*

At this stage of the question and with scanty data, it is unwise to make any conclusive serious statement about the typological absence of certain syncretisms. However, one may advance at least one: even though the syncretism comitative–instrument is common cross-linguistically for instrumental cases, the comitative is

not found in ergative syncretisms that encode instrument or otherwise (Stolz 1996b, Palancar 2002).[3]

37.2.2 The origin of ergatives

What do these syncretic patterns reveal about ergative case? Ergativity has long been seen in linguistics as an emergent structure that came about through the reanalysis of a nominative–accusative alignment. In this view, passive has been believed to be the most common path for this reanalysis, especially for the Indo-Iranian, Polynesian, and Eskimo ergative languages (for a full discussion on the matter see Palancar 2002). However, the evidence provided is not fully convincing, and there are many who remain sceptical. Recently, Butt and King (1991, 2003, 2004) have challenged such a view with convincing arguments, for example the old Sanskrit inflectional instrumental with multiple allomorphy used with demoted agents in passive-like structures could not possibly have given rise to current ergative markers, as these are of recent development, and behave like clitics. Butt and King's solution is that ergative case *=ne* in Urdu/Hindi stems from a grammaticalized adposition *janiyē* 'for the sake of, because of', which was used in the past tense to emphasize semantic agentivity. All in all, ergative syncretisms are often taken to shed light on the possible emergence of the alignment they manifest. Palancar (2002) treats ergative syncretisms involving instrument (and cause) as reflecting a more abstract energetic category he calls 'energiser', much along the lines of the 'effector' category in Van Valin and Wilkins (1997). Such a view proves more convenient than explaining which was first, instrument or agent. The oblique ergatives of the Caucasian type reflect very old markers whose history is difficult to trace. As for the spatial cases in the syncretisms, the spatial categories involved (mainly source and location) served as a model to metaphorically construe other more abstract, semantic categories such as agent/instrument/cause in order to foreground the energetic role played by such participants in the clausal event.

Abbreviations (languages, language groups, and regions)

AT Australian
AU Austronesian
CA Caucasian

[3] The apparent exception to this rule is found in Circassian languages. As the ergatives in these languages are really oblique markers, an NP functioning as comitative can also receive this case providing the relevant cross-referencing morphology appears in the verb. I thank Yuri Lander for this observation.

978–0–19–920647–6 38-Spencer-ch37 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 571 of 571 May 16, 2008 19:0

E	East
IE	Indo-European
LI	Language Isolate
MJ	Macro-Jê
NPN	Non-Pama-Nyungan
NW	North West
P	Papuan
Pan	Panoan
Pol	Polinesian
TB	Tibeto-Burman
TNG	Trans-New-Guinea
UY	Uralo-Yukaghir
W	West

CHAPTER 38

VARIETIES OF DATIVE

ÅSHILD NÆSS

38.1 Problems of definition

Of the commonly recognized case categories, the dative is perhaps the most difficult to define in a consistent, cross-linguistically valid way. In Greek and Latin, where the term originates, the dative case was used both for the goal/recipient argument of ditransitive verbs, and for the complements of certain intransitive verbs such as 'help', 'obey', or 'trust'. These two types of argument are commonly taken to share the property of being 'indirectly affected' by the verbal event, and so given the label 'indirect objects'. However, not all languages use the same formal marker for these two functions, thus raising the question of which of the two, in languages where they are formally distinct, should be counted as the 'dative'.

Blake (2001) argues that in such instances, the 'recipient' case should be singled out as distinct, while the case of the complements of verbs similar to those cited for Greek and Latin should be labelled dative. He thus considers the dative 'the main noncore case used to mark complements' (Blake 2001: 143). However, many others in practice appear to equate the dative with the case of the recipient of ditransitives (e.g. Blansitt 1988, Van Belle and Van Langendonck 1996). The problem may have no obvious solution: 'In the end, one has to acknowledge that the basis for naming a case "Dative" in a particular language is subject to some variability' (Newman 1998: 11).

978–0–19–920647–6 39-Spencer-ch38 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 573 of 580 May 16, 2008 19:0

A related difficulty concerns the place of the dative in relation to the commonly assumed dichotomy between 'structural' and 'semantic' cases. The dative almost by definition appears to straddle the structural–semantic divide; as the case of ditransitive indirect objects, it may be said to have a structurally specifiable function, but it also tends to have uses which can only be accounted for in semantic terms, such as marking experiencers or beneficiaries (cf. (1) below).

These problems of definition naturally lead to problems of delineation with respect to other case categories particularly local cases, as markers covering the typically 'dative' functions also frequently have locative uses (see section 38.5 and Creissels, Chapter 42). In such cases, which label to apply to the marker in question may ultimately come down to a matter of perspective. For instance, the case suffix labelled 'allative' in West Greenlandic (Fortescue 1984) covers many of the functions typically associated with the dative, such as recipient of ditransitives, benefactive, and purpose, as well as spatial goal and location. In languages where case relations are indicated by adpositions rather than inflectional morphology, this distinction may be particularly difficult to make; the 'locative-directional' preposition *i* in Samoan, for example, has a wide range of non-spatial, dative-like uses (Mosel and Hovdhaugen 1992).

38.2 Core meaning of the dative

As will be clear from the introduction, case-markers labelled 'dative' vary greatly in the range of functions they cover, although they must be assumed to share some common semantic core. The uses which might be said to define the category 'dative' as it is commonly understood are to mark recipients, benefactives/malefactives, experiencers, goals, and purposes. Lezgian is an example of a language where the dative covers most of these central functions:

(1) Lezgian (Nakh-Daghestanian; Haspelmath 1993a: 88–9):

a. *Ruš.a* ***gada.di-z*** *cük* *ga-na*
girl(ERG) boy-DAT flower give-AOR
'The girl gave a flower to the boy.'

b. *Anglija.d-a* ***lif.re-z*** *gümbet* *ecig-nawa*
England-INESS dove-DAT gravestone build-PRF
'In England a gravestone has been built for a dove.'

c. ***Kasbuba.di-z*** *tar.a-n* *xil.e-l* *zurba* *sa* *q̃uš* *aku-na*
K.-DAT tree-GEN branch-SRESS big one bird see-AOR
'Kasbuba saw a big bird on a tree's branch.'

d. *Ča-z muhman-ar ata-na*
we-DAT guest-PL come-AOR
'Guests came to our place.'

In addition, the dative is frequently found with objects of intended or uncompleted actions (2a), objects which are low in affectedness (2b–c) or animate as opposed to inanimate objects (2c–d):

(2) Waris (Trans-New Guinea; Brown 1988)
a. *Ungevli-rini téh-m ga-v*
women-ABL firewood-DAT go-PRES
'A woman is going (to get) firewood.'
b. *Ka-va ti-m he-the-v*
1-TOP tree-DAT chop.down-PROCESS-PRES
'I am chopping on a tree.'
c. *Ka-va ti he-v.*
1-TOP tree chop.down-PRES
'I am chopping down a tree.'
d. *Ka-va ye-m hélvakomandha-v*
1-TOP 2-DAT kill-PRES
'I kill you.'

Clearly, these senses are not all completely distinct semantically. Two generalizations immediately suggest themselves. Firstly, leaving out ditransitives for the time being, the dative case tends to be associated with **relatively low transitivity**: with verbs that only indirectly affect their objects, such as 'help' or 'obey'; with actions which are not yet completed and so have not yet produced any perceptible effects; or with atypical subject participants such as experiencers. Secondly, the dative frequently reflects the affectedness, direct or indirect, of **animate or sentient** entities – this includes the recipients of ditransitives, which must typically be animate in order to be plausibly said to receive something, as well as experiencers, benefactives/malefactives, and animate objects.

These two broad generalizations are themselves not unrelated. Firstly, sentience or volitionality on the part of an affected participant often makes the participant a less typical patient, which may lead to a lower degree of semantic transitivity. Prototypical patient-objects do not exercise any degree of control over the action which affects them, and the presence of such control often implies a less radical effect on the participant which allows itself to be affected. In Icelandic, *Hann klóraði mig* 'He scratched me' with an accusative object refers to scratching as an act of violence, probably causing the patient pain, whereas *Hann klóraði mér* with a dative object means 'He scratched me (to relieve an itch)' (Barðdal 2001: 146). The dative-marked object is understood as voluntarily submitting to the scratching, which is consequently interpreted as less severe than when the object is a genuine patient affected against his/her will. Note also that such a volitionally involved patient is

almost indistinguishable semantically from a beneficiary; the most obvious motivation for volitionally submitting to an act is that one expects its effect to be beneficial.

Secondly, the range of potential 'indirect' effects is clearly far wider in the case of animate/sentient participants – receiving, benefiting from, or experiencing something presupposes the sentience of the participant in question, while inanimate participants do not easily register any effects beyond direct physical impact. That is, dative marking typically implies not just affectedness, but a particular **kind** of effect, namely an effect which can only be ascribed to sentient participants. It should be added that the sentient, affected participant is generally not seen as actively controlling the event which affects it, that is, a dative-marked participant is typically distinct from what Saksena (1980) calls an 'affected agent'.

There is clearly no necessary connection between the affectedness of an animate entity and low transitivity – on the contrary, examples like (2d) refer to what must be understood as a highly transitive situation. But the notion of animate affected entities as being less prototypical patients by virtue of their sentience can be taken to provide a semantic link between these different, seemingly unrelated uses of the dative.

This understanding of the core meaning of the dative clearly implies that dative marking most typically occurs on animate rather than inanimate nouns. However, the fact that dative markers frequently also have spatial uses (see section 38.5 below) means that such markers do in many languages also appear on inanimate-referring NPs. There has been considerable debate concerning which function of the dative should be taken to be primary – the local functions which typically include goal, source, and location, or the use to mark various types of sentient affected entities, as decribed above. For the Indo-European dative, evidence for one or the other function as historically primary is conflicting (Van Langendonck 1998: 253). However, a case with purely or predominantly local functions is perhaps better labelled e.g. allative or locative rather than dative.

It is not unusual for a language to have a distinct inflectional dative in pronouns but not in nouns, as in French *il lui donne le livre* 'he gives him the book' vs. *il donne le livre à Jean* 'he gives the book to Jean'. It is also possible for a language to formally mark dative-type nominals but make no nominative–accusative distinction on nouns; this is the case for example in Iatmul, where the 'allative' suffix is used for recipients and benefactives as well as spatial goals; thus there is no case-marking on the A and O arguments, but overt marking on the recipient in examples like (3):

(3) Iatmul (Sepik; Foley 1986: 96):
ntɨw waalə wɨ-ŋkət kwɨy-ntɨ
man dog I-ALL give-3SG.M
'The man gave me the dog.'

Similarly, it could be argued that English, which marks dative-like functions by means of prepositions (*to/for the man*), has a dative but no accusative for lexical nouns.

The peculiar status of the dative as a partly 'structural', partly 'semantic' case may be understood as deriving from its core meaning of 'sentient affected entity'. On the assumption that all cases have semantic content, the 'structural' cases are those whose core semantics allow them to be generalized to the point where their distribution can be given a purely structural definition. For instance, if the core meaning of the accusative is marking the patient of a transitive event, it is a fairly small step to generalize the use of this case to all objects of formally transitive clauses, regardless of whether they are, strictly speaking, patients.

Such generalization is possible for the dative only in the subset of contexts where its use can be related to a structural position, namely that of indirect object of ditransitives. Participants undergoing the particular kind of sentience-conditioned affectedness implied by the dative occur in a range of different constructions, only some of which – recipients or beneficiaries of formally ditransitive verbs – readily lend themselves to description in structural terms. The remainder require a semantic description at least to some extent, with the result that the dative is perceived as having both structural and semantic functions.

38.3 Alternative marking strategies

It is not uncommon for dative-like functions to be expressed through formal means other than bound nominal morphology. A number of languages have verbal cross-referencing affixes which index just the types of participants which in other languages are marked by dative case on the NP (see for instance Wichmann, Chapter 56, on Tlapanec). An example is Koasati, where the 'position 5' verb prefixes refer to typical dative functions such as indirect objects of ditransitives, benefactives, subjects of verbs such as 'be hurt' and 'be hungry', and objects of 'indirect-effect' verbs such as 'help', 'obey', and 'amuse':

(4) Koasati (Muskogean; Kimball 1991: 131–2):

a. *st-am-íl*
INS-1SG.DAT-arrive.here
'Bring it to me!'

b. *cim-acó:li-t*
2SG.DAT-sew-PAST
'She sewed it for you.'

c. *kom-akásn*
1PL.DAT-be.hungry
'We are hungry.'

d. *cim-há:lo-li-laho-V*
2SG.DAT-hear/obey-1SS-IRR-PHR:TERM
'I will obey you.'

In some languages, participants in typical dative functions are cross-referenced with the same markers otherwise used for the objects of transitives. That is, only two participants are cross-referenced on the verb at any one time, and if there are three participants in the clause, it is the dative rather than the accusative which is cross-referenced. This is the case for example in Huichol (Uto-Aztecan), where the 'object agreement prefix' cross-references the object of a monotransitive verb, but the recipient or benefactive participant if such a participant is present in the clause. In the latter case, the (patient/theme) object is not cross-referenced; thus *meci* '2SG' cross-references the transitive object in *taame eek ɨ te-meci-zeiya* (we you 1PL-2SG-see) 'we see you', but the recipient in *nee waakanaari ne-meci-t ɨkiit ɨ eek ɨ* (I chickens 1SG-2SG-give you) 'I gave the chickens to you' (Comrie 1982: 99, 107; for more examples see Dryer 1986).

It is fairly common for dative functions to be indicated by adpositions, as in English and many other languages of Europe. This is frequently the same adposition used for spatial goals or directions, cf. sections 38.5, 38.7 below.

38.4 Functional varieties

As the core uses of the dative listed above are fairly diverse, it is to be expected that not all languages with a 'dative' employ the same marker for all these uses; the semantic domain generally subsumed under 'dative' is subdivided in different ways in different languages. Distinct markers for purposes and/or benefactives are not uncommon, and some languages group recipients with spatial goals rather than with other types of 'indirectly affected' participants (see below).

A distinct purposive case is found e.g. in Yidiny (*-nda* 'DAT' vs. *-gu* 'PURP') and Irula (*-(u)kku/-kke* 'DAT' vs. *-kk/-kkāyi* 'PURP'). Kalkatungu has one marker (*-kunha*) for recipients and spatial goals, another (*-ku*) for purposes, benefactives, and complements of certain intransitives; Blake (2001) labels the former allative and the latter dative. Basque has both a benefactive and a purposive case, in addition to the dative which is used on indirect objects of ditransitives, experiencers, and animate goals of motion, among other things. Warrgamay uses the genitive for recipients of verbs like 'give', while the dative is used for purposes and 'indirect objects' of verbs such as 'wait for', 'sing for' (Dixon 1980).

Many languages which have a dative case use some other form of marking for at least some experiencer arguments. This may be due to the fact that the label 'experiencer', as it is typically used, refers to a semantically rather heterogeneous category. For example, it is common for languages to use the canonical transitive

patterns for experiencers which are perceived as being to some extent in control of the experienced event, while the dative is used for experiencers which are not ascribed such control. Godoberi has a distinct 'affective' case, marked by *-ra*, which is only used for the experiencers of the verbs 'see', 'hear', 'know', and, optionally, 'forget'; other experiencers take the dative marker *-łi* which also covers recipients, beneficiaries, and purposes (Kibrik 1996: 79–80).

38.5 Overlap with other cases

Given the fairly complex semantic domain typically covered by the dative case, it is not surprising that formal markers of dative-type functions often also have uses associated with other case categories.

Dative–accusative syncretism commonly occurs in languages with differential marking of direct objects on the basis of definiteness or animacy (see Kittilä and Malchukov, Chapter 36); in such languages, the case-marker used on definite/animate objects is often formally identical to the dative (cf. example 2c–d). This type of syncretism is generally assumed to be linked to the greater degree of independence or individuation of such objects, which would make them similar to the type of participant typically marked by the dative.

This analysis presupposes that it is the original dative case which gets extended to accusative function, as indeed is very frequently the case, e.g. in many Romance and Indo-Aryan languages. The opposite appears to occur in Yukaghir, where the accusative can be used for recipients and addressees, functions otherwise covered by the dative (Maslova 2003). Similar extensions of the accusative to some dative functions can be found for example in some Dravidian languages (Krishnamurti 2003).

Dative–allative syncretism is similarly very frequent, probably explainable in terms of the semantic similarity between a recipient or other 'target' of an event and a spatial goal. *Dative–locative* syncretism also occurs; Blansitt (1988) argues that a single marker can only be used to cover both dative and locative function if it also covers the allative.

Dative–genitive syncretism occurs in a number of varieties. A number of Australian languages have a single form covering both dative and genitive functions. Some languages, e.g. many Dravidian languages, employ the dative for a restricted set of possessive relations, typically kinship and body-part relations. There are also languages where the genitive extends to cover certain dative-like functions, such as Bengali where it is used for some experiencers, and Warrgamay where it marks recipients.

38.6 Other uses

A number of languages employ the dative for the causee argument of causative constructions. In some languages, for example Turkish, this only applies to causativized transitives, and so the use of the dative here may be seen simply as an extension of the dative as a marker of indirect object to include derived ditransitives. In other languages, however, the dative may also be used on the causee of intransitives, and may sometimes alternate with another case, with a resulting difference in meaning. Typically, the dative contrasts with the accusative in such instances, with the dative-marked causee interpreted as having a greater degree of control over the event than an accusative-marked causee. Japanese is a well-known example of a language showing this alternation:

(5) Japanese (Isolate; Comrie 1985: 334)
   a. *Taroo ga Ziroo o ik-ase-ta*
      T. NOM Z. ACC go-CAUS-PAST
      'Taroo made Ziroo go (forced him).'
   b. *Taroo ga Ziroo ni ik-ase-ta*
      T. NOM Z. DAT go-CAUS-PAST
      'Taroo made Ziroo go (persuaded him).'

The dative is also found as an oblique marker in valency-decreasing constructions. Most commonly, it marks the demoted object of antipassives, for example in a number of Australian languages; but there are also languages where the dative is used for passive agents, e.g. Korean:

(6) Korean (isolate; Sohn 1994: 231)
   a. *Minca-ka ton-ul apeci-hanthey tuli-ess-ta*
      M.-NOM money-ACC father-DAT give-PAST-DECL
      'Minca gave the money to her father.'
   b. *Minca-nun Yonggo-hanthey mac-ass-ta*
      M-TC Y.-DAT beaten-PAST-DECL
      'Minca was beaten by Yongho.'

In some languages, the dative has functions related to modality or evidentiality; this is the case for example for the Georgian 'inversion' construction with dative subject: *Turme Rezo-s samajuri učukebia dedastvis* (apparently Rezo-DAT bracelet he.gave.it.EVID mother.for) 'Apparently Rezo gave the bracelet to Mother' (Harris 1984: 259). Pitta-Pitta uses an antipassive construction with a dative object to indicate desiderative modality; Yukulta uses the same construction in irrealis contexts. In Latin gerundive constructions, for example, and with certain verbs in Hindi-Urdu, dative marking on subjects may indicate obligation (e.g. Hindi-Urdu *Anjum-ne/-ko xat lik^{h}-naa hai* (Anjum-ERG/-DAT letter write-INF is) 'Anjum wants

to (ERG)/has to (DAT) write a letter'; Butt and King 1991: 34). The latter use might be seen as related to the use of the dative on subjects which act involuntarily or are not perceived to be in control of their act, occurring e.g. in German *Mir ist der Teller zerbrochen* (I.DAT is the.NOM plate broken) 'I accidentally broke the plate; the plate broke on me'.

In some languages, a marker formally identical to the dative occurs on verbs. Most commonly cited examples are from Australian languages, where the dative occurs as a purposive marker on verbs in e.g. Warrgamay and Kalkatungu, and as a marker of incompletive aspect in Kala Lagaw Ya (Blake 2001: 180–2). The Kayardild 'verbal dative' marks typical dative functions such as recipient, beneficiary, and goal, but has the peculiarity (along with five other 'verbal cases') of converting the argument to which it attaches into a morphological verb:

(7) Kayardild (Tangkic; Evans 1995a: 164)
niya *waa-jarra* *wangarr-ina* *ngumban-maru-tharra*
3SG.NOM sing-PAST song-MABL your-VDAT-PAST
thabuju-maru-tharr?
elder.brother-VDAT-PAST
'Did she sing the song for/to your elder brother?'

38.7 Diachronic issues

A common diachronic source for datives is from an allative or other marker of direction, as in e.g. Turkish, Quechua, or the Romance languages; some languages mark dative-type participants with a marker deriving from the verb 'come' or 'go' (Newman 1998). In a number of languages, the dative marker has developed from a verb meaning 'give', for instance in Mandarin, Thai, and a number of Niger–Congo languages (Givón 1975, Blake 2001). Dative markers commonly develop further into accusatives, cf. section 38.5 above.

CHAPTER 39

VARIETIES OF GENITIVE

YURY A. LANDER

39.1 From function to form

39.1.1 Core function

Genitive is basically an adnominal case and in fact, a basic adnominal case. Its core function is to mark a nominal whose referent (**possessor**) is connected by a **possessive relation** to the individual expressed by the phrase within which the genitive phrase is embedded (**possessum**). Hence genitive constructions are a kind of possessive construction (**possessive**) and any discussion of genitives requires a discussion of possessives in general.

The possessive is the most unmarked adnominal construction expressing a relation. This is reflected by the remarkable frequency of possessives, their wide distribution, not uncommon formal simplicity (as compared to other constructions of this kind) and most importantly, semantic unmarkedness. Possessive relations are usually induced from the context and the lexical semantics of the corresponding nominals (Partee 1997; Barker 1995). Thus, (1) could mean 'the book that Liza owns' or 'the book that Liza wrote' or 'the book that Liza photographed' etc., depending on the context:

978–0–19–920647–6 40-Spencer-ch39 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 582 of 592 May 16, 2008 19:0

(1) Russian
knig-a *Liz-y*
book-NOM.SG Liza-GEN.SG
'Liza's book'

Nonetheless, some relations arise in certain semantic configurations by default, the most prominent of which is the relation of control/ownership. Furthermore, some languages (especially in Oceania) have relational classifiers (Lichtenberk 1983; Aikhenvald 2000), which restrict the possessive relation, while in other languages the possessive relation is sometimes specified by means of a modifier (see Ackerman 1998 and Partee and Borschev 2000 for discussion). For example, in (2), the relation ('being stolen by') between the possessum ('money') and the possessor ('I') is specified by a participle:

(2) West Armenian (Ackerman 1998)
(*im*) *koʁts-adz* *təram-əs*
I.GEN steal-PTCP.PAST money-1SG
'the money I stole'

Prototypical possessives establish the reference of the possessum via the possessor, an 'anchor' which links the possessum to the context (Keenan 1974; Langacker 1993, 1995; Taylor 1996). This leads to the tendency of (more prototypical) possessives to be specific/definite, which may result in their occasional incompatibility with determiners (Lyons 1986; Haspelmath 1999b) or their use in the expression of contextual definiteness (Fraurud 2001). The best anchors are topical (concrete, animate, definite, and, in the best case, pronominal), hence the inclination of possessors to be more topical. Constructions with non-topical possessors, as in (3), are non-prototypical for possessives; such meanings are often conveyed by non-possessive constructions (e.g. with adjectives).

(3) Georgian (Koptjevskaja-Tamm 2004: 159)
p'ur-is *dana*
bread-GEN knife
'a/the bread knife'

39.1.2 Marking possessives

Since genitive marking is not uniform, it is convenient to begin its description with an excursion into formal types of possessives in general. Possessives utilize various means of marking, which can be classified according to numerous parameters (see Ultan 1978; Seiler 1983; Nichols 1988, 1992; Croft 1990: 27–38; Plank 1995b: 38ff.; Koptjevskaja-Tamm 2003a), of which I will discuss three, namely the locus of marking, indexing, and the degree of synthesis.

978–0–19–920647–6 40-Spencer-ch39 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 583 of 592 May 16, 2008 19:0

The most widely used typology of possessive marking is concerned with its locus, that is the participant associated morphosyntactically with this marking. A possessive marker can be associated with the possessor (dependent-marking, as in 1–3) or with the possessum (head-marking, as in 4), or with both (neutral marking, example 5):

(4) Apalai (Koehn and Koehn 1986: 85)
nohpo kyry-ry
woman thing-POSS
'the woman's possession'

(5) Eastern Pomo (McLendon 1975: 167)
ká'wk-i-Yà
person-POSS-bone
'human bones'

The asymmetry between the participants can also be reflected by word order only (Malay *buku Umar* 'Umar's book'), or by incorporation of the possessor into the possessum (Egyptian Arabic *kitaab-muna* 'Mona's book', Gary and Gamal-Eldin 1982: 48).

Various sorts of marking can combine within one construction. For example, in (2) above, the head-marking suffix cross-referencing the possessor coexists with the dependent-marking genitive.

Next, the possessive marking can be divided into two distinct categories, one that indexes features of a participant other than its locus, and one that does not do so. A comparison of (4) with (2) illustrates this distinction in head-marking constructions. An example of an indexing dependent-marking construction follows (numerals designate noun classes):

(6) Rwanda (Dubnova 1984: 44)
iki-bindi cy'-umu-gore
VII-vessel VII.ATTR-I-woman
'(the) woman's vessel'

Again, non-indexing marking can combine with indexing. Where the latter is based on case concord, we can find two case markers on one nominal, the phenomenon called ***Suffixaufnahme*** (Plank 1995a):

(7) Awngi (Hetzron 1995: 326)
wolijí-w-des aqí-w-des ŋɔ́n-des
old-GEN-ABL man-GEN-ABL house-ABL
'from the old man's house'

Finally, possessive marking can be classified into synthetic and analytic. Thus, instead of affixes broadly illustrated above (and suprasegmental means like tone, as

978–0–19–920647–6 40-Spencer-ch39 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 584 of 592 May 16, 2008 19:0

in Burmese), dependent-marked possessives frequently involve adpositions, while head-marking possessives may utilize clitics. A kind of possessive close to analytic head-marking is found in some Germanic idioms:

(8) Colloquial High German (Koptjevskaja-Tamm 2003a: 666)
mein-em Vater sein Buch
my-DAT.M.SG father his book
'my father's book'

Occasionally, characterization of a construction according to the above-mentioned parameters (head-marking/dependent-marking, indexing/non-indexing, and synthetic/analytic) can be problematic. For example, the construction in (8) may display characteristics of dependent-marking and even neutral-marking, together with head-marking ones (Koptjevskaja-Tamm 2003a; Strunk 2004). Still, for most data these parameters remain convenient classificatory tools.

39.1.3 Formal expressions of genitive

Given the multiplicity of means of marking found in possessives, a question can be asked: what marking is used in genitive constructions? It seems, however, that the only restrictions on this marking are based on the case status of genitive.

Being a case, genitive by definition constitutes a dependent-marking strategy which is paradigmatically contrasted with marking non-possessive relations. The synthetic genitive is most often expressed with suffixes. However, in Mangarayi the genitive/dative case is sometimes marked by a prefix (e.g. *ŋaya-gaḍugu* 'GEN-woman'; Merlan 1989: 57), and in the extinct Mochica the contrast between the nominative and the genitive forms of singular personal pronouns was conveyed by vowel alternation:

(9) Mochica (Adelaar with Muysken 2004: 331)

	Nominative	Genitive
1SG	moiñ	mœiñ
2SG	tzhang/tzha	tzhœng

While synthetic genitives are easily recognized, analytic genitives are often treated as adpositions. Where such markers stand in a paradigmatic contrast to similar role markers, they should be considered case expressions. In Tukang Besi the genitive marker apparently attaches to a constituent and is usually contrasted with other cases:

(10) Tukang Besi (Donohue 1999: 344)
te lai u Sentani kene Kota raja
TOP distance GEN Sentani and Kota raja
'the distance between (lit. of) Sentani and Kota raja'

Both synthetic and analytic genitive markers may simultaneously express other categories. Most commonly, this involves number (as in Latin *puer-i* 'boy-GEN.SG' vs. *puer-orum* 'boy-GEN.PL'). Another pattern is observed in German, where the genitive is sometimes found only in determiners which simultaneously mark definiteness (cf. *das Bild ein-er Frau* 'the.NOM.N.SG picture a.GEN.F.SG woman'). Note, however, that definiteness/specificity can also be *implied* by the very presence of genitive marking (see below).

A particular pattern to be distinguished from the simultaneous expression of genitive with other categories is reported for Tsez (and some closely related languages), where the choice of a genitive marker depends on the case of the possessum: the direct genitive is found with the absolutive possessum (11a), while the oblique genitive occurs with the possessum in other cases (11b). This is a rare instance of an indexing non-concord genitive.

(11) Tsez (Kibrik 1995: 222)

a. *obi-s* *esij* *idu-r* *ajsi*
father-GEN.DIR brother[ABS] house-INESS came
'Father's brother came home.'

b. *obi-z* *esi-s* *joł* *ʁIutku*
father-GEN.OBL brother-GEN.DIR be house[ABS]
'Father's brother owns a house.'

Languages that display multiple strategies of genitive marking are not rare. Usually they divide between pronominal and non-pronominal possessors, where the former have suppletive or more grammaticalized forms. Sometimes pronominal possessors have a distinct position, as in Russian, where genitives usually follow the possessum but third person pronominal genitives typically occur before their heads (cf. *ego otec* 'he.GEN father[NOM.SG]').

Thus, genitives can exploit a number of marking strategies, and the main limitations on them relate to the very definition of genitive as a case.

39.1.4 Restrictions on the use of genitive

Genitive constructions often contrast with possessives of different kinds, which occasionally restricts the use of a genitive marker to certain classes of nominals. There are two tendencies manifested in such restrictions. The first relates to the fact that the genitive is inflection and as such is prone to mark more topical nominals, for which more inflectional distinctions are normally found. A revealing example of a split resulting from this is found in Sinhala, where, judging from Vyxuxolev 1964, the genitive is used with animate possessors only, while inanimate possessors choose the locative. A similar situation is found in a subtype of possessives expressing part–whole relations in some Pama-Nyungan languages and in Imbabura Quechua,

where animate possessors require the genitive (12a), while genitive marking on inanimates (12b) is odd:

(12) Imbabura Quechua (Cole 1985: 117)
a. *alku-paj uma*
dog-GEN head
'the head of the dog'
b. *yura(ʔ-paj) uma*
tree(-GEN) head
'the top of the tree'

That the source of such splits is not animacy but topicality is confirmed by languages where the presence of the genitive depends on other topicality features. For example, in Turkish the genitive suffix appears on specific possessors only:

(13) Turkish (Maizel 1957: 17)
a. *kadın-ın şapka-sı*
woman-GEN hat-3SG
'some/the woman's hat'
b. *kadın şapka-sı*
woman hat-3SG
'a women's hat'

The second tendency, which partly resists the first, is related to the fact that genitive is opposed to many other possessive means in that it can supply possessors with greater syntactic autonomy. Consequently, at times genitive is not used with more topical possessors, which often function as anchors only and do not need any autonomy. Thus, in Bagvalal, more topical pronominal, singular masculine, and plural human possessors (plus some toponyms) show concord with the possessum (14a), while others employ the genitive (14b):

(14) Bagvalal (Daniel 2001: 140)
a. *ehun-dar-alu-b misa*
blacksmith-PL-OBL.HUM.PL-NHUM house
'the blacksmiths' house'
b. *χ̣an-ē-Ī un-abi*
horse-OBL.PL-GEN head-Pl
'the horses' heads'

Both tendencies manifest themselves in Yiddish: pronominal possessors show concord with the possessum, and human (or personified) nominal possessors are usually marked with the genitive case, while less topical possessors are introduced by a preposition.

978–0–19–920647–6 40-Spencer-ch39 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 587 of 592 May 16, 2008 19:0

39.2 Functional varieties and polysemy patterns

39.2.1 Functional varieties

Possessives show various oppositions depending on certain characteristics of the possessor, the possessum, and the possessive relation. However, such oppositions are only very rarely conveyed by contrasting different genitive exponents. Instead, functional splits are usually expressed by oppositions between the genitive and other constructions or reflected in variation in syntactic characteristics of genitive phrases.

The possessum's type can affect the choice of a possessive in that some languages distinguish between definite and indefinite possessives. Genitive encoding typically remains neutral to this, though in some languages (e.g. in German) the position of a genitive phrase can imply definiteness. Furthermore, syntactic characteristics of those possessors that do not determine the reference of the possessum but only serve as restrictive modifiers (such as non-specific possessors and arguments of so-called 'picture nouns' like *Lisa Gherardini* in *a/the portrait of Lisa Gherardini*) are often different from possessors serving as full-fledged anchors in their position and capacity to coordinate with other modifiers, be relativized, stacked, etc. (see Koptjevskaja-Tamm 2004; Grashchenkov 2006 among others).

Occasionally it matters whether the possessive relation is given (more or less) unambiguously: thus, **relational nouns** (such as *brother* or *top*) presupposing the existence of some possessor exhibit special behaviour in comparison with other nouns. The distinction between relational and non-relational nouns, when grammaticalized, is called the 'inalienable/alienable distinction' (see Chappell and McGregor 1995). Nichols (1988; 1992) argues that in languages having this opposition, dependent-marking, including genitive, normally marks alienables. However, this does not hold for languages, such as Kuot (Lindström 2002), Budukh, Khinalug, and possibly Polynesian languages, which convey the alienability distinction solely by means of dependent-marking. Khinalug even distinguishes between alienable and inalienable genitives for some nouns (although surprisingly, relational kinship terms belong to the alienable class here):

(15) Khinalug (Kibrik et al. 1972: 131–2)
  a. *gad-i* *c'u*
     boy-GEN.INAL name
     'the boy's name'
  b. *gad-e* *bij*
     boy-GEN.ALIEN father
     'the boy's father'

Notably, inalienable possessors also can have specific syntactic properties. For example, in Japanese, only inalienable possessors allow relativization (Iwasaki 2002: 181).

Finally, it is sometimes important whether the possessor is used exclusively for restricting the reference of the possessum or is relevant to the whole utterance (Lander 2004). This is often reflected by the distinction between adnominal and external (clause-level) possessives (Payne and Barshi 1999). External possessors typically employ a case other than genitive (16), but sometimes genitive appears at the clause-level (17). Furthermore, many languages allow the use of a genitive phrase without the possessum (18).

(16) French (Lamiroy and Delbecque 1998)
On lui a cassé le pied
one he.DAT has broken the foot
'They broke his foot.'

(17) Udi
χačan-i=jal sa eʁel=e bak-sa
Khachan-GEN=and one ram=3SG be-PRES
'And Khachan has a ram.'

(18) Basque (Saltarelli 1988: 161)
diru-a ama-ren-a d-a
money-ABS.SG mother-GEN.SG-ABS.SG 3.ABS-PRES(-be)
'The money is mother's.'

39.2.2 Non-possessive uses of genitive

At the phrase-level, genitive sometimes appears to mark dependents of the heads that do not describe entities. The most frequent patterns of this kind are described below.

(i) It is not unusual for genitives to mark arguments in constructions such as (19), where a verb acquires nominal features in order to serve as an argument of some other predicate.

(19) Lithuanian (Koptjevskaja-Tamm 2003b: 733)
Kolumb-o Amerik-os atradi-ma-s
Columbus-GEN.SG America-GEN.SG discover-AN-NOM.SG
'Columbus' discovery of America'

Note, however, that cross-linguistically this is just one of many ways of coding such dependents, and languages vary considerably as to which arguments can be coded by genitive (Koptjevskaja-Tamm 1993; 2003b).

978–0–19–920647–6 40-Spencer-ch39 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 589 of 592 May 16, 2008 19:0

(ii) In a number of languages like Turkish, Manchu, Chukchi, Sinhala, and Japanese, genitive can code the subject of a relative clause:

(20) Japanese (Andrews 1985b: 27)
kore wa ano hito ga/no kai-ta hon desu
this TOP that person NOM/GEN write-PAST book is
'This is the book which that person wrote.'

Sometimes this can be explained by the participle use of nominalizations, yet in other cases the source of this function of genitive can be a construction such as (2) specifying the possessive relation.

(iii) Occasionally genitive is used in quantificational constructions, presumably due to the nominal origin of the relevant quantificational elements, or to the fact that such items frequently have some properties of the nominal head (cf. Abney 1987). The following example illustrates genitives governed by a numeral and a measure noun:

(21) Ossetic (Axvlediani 1963: 169)
æxsæz xos-y t'yfyl-y
six hay-GEN pile-GEN
'six piles of hay'

(iv) In a few languages genitive participates in constructions involving several nominals characterizing the same referent.

(22) Bagvalal (Daniel 2001: 150)
mosku(-ɬ) mak'a
Moscow(-Gen) place
'the city of Moscow'

For other kinds of possessives, a related phenomenon occurs whereby a semantic modifier appears as the head while the noun determining the reference of the phrase is marked as a possessor (Ross 1998; Malchukov 2000).

(v) Another frequent occurrence of genitive concerns marking of the objects of adpositions and spatial adverbials:

(23) Aghul (Merdanova 2004: 29)
Xul-ar-i-n üdih
house-PL-OBL-GEN front
'in front of the houses'

Many such heads in fact go back to nouns, which motivates possessive marking of their dependents.

At the clause level we find a number of instantiations of genitive that are *prima facie* unrelated to the 'possessive genitive'. Some of them indicate the reduced individuation of nominals, the classic example being the 'genitive of negation' found

in some Slavic and Germanic languages. Here the choice of the genitive rather than some other case can be motivated by the non-specificity of a nominal under the scope of negation (24). Parallels to this are observed, for instance, in Kuot, where the possessive marker may also accompany deindividuated core participants.

(24) Russian
Nikogda on ran'še mandarin-ov ne vid-yva-l.
never he.NOM earlier tangerine-GEN.PL not see-ITER-PAST[M.SG]
'He had never seen tangerines before.'

Furthermore, genitive marking of either actor or undergoer can reflect reduced semantic transitivity (in the sense of Hopper and Thompson 1980). For example, in some Latin strata we find object genitive with low transitive verbs such as *studeō* 'to strive after', and *fastidiō* 'to dislike'. This function is possibly related to the preceding one, since deindividuation of a participant often goes together with reduction in transitivity.

Thus, clause-level genitive can code marked situations as compared to more prototypical instances of the relevant constructions licensing other cases. However, we will see in the next section that such markedness is by no means universal.

39.2.3 Polysemy patterns

It is not uncommon for a language to utilize the same case for the coding of adnominal possessors and for some basic function at the clause level, though such a case is often not labelled as 'genitive'. Two kinds of polyfunctionality can be distinguished in this respect. The first kind includes polysemy with core/syntactic cases. For example, languages often code the possessor in a similar way to the marked participant in a transitive construction. Thus, possessives reflect ergative case in such diverse languages as Eskimo, Austronesian Niue, Indo-European Ladakhi, Northeast Caucasian Lak, Northwest Caucasian Circassian languages, and an isolate, Burushaski (though in Circassian and Burushaski it may be better to speak of a general oblique case). There are also languages such as Martuthunira and Karachai-Balkar where the possessor is marked identically to the transitive object. Finally, in many Philippine languages claimed to belong to the symmetric type (Foley 1998), one case codes both the possessor and the non-subject argument in transitive clauses (i.e. undergoer in active transitives and actor in passive transitives).

The second kind of polyfunctionality concerns more semantic cases such as dative, ablative, and locative. Quite often, possessors are marked like recipient/beneficiary datives. This is observed, for instance, in Bulgarian, Colloquial French, languages of Australia, etc. That the connection between the two functions is not accidental is supported by the fact that even in languages contrasting datives and genitives, the former occasionally are found on adnominal-anchoring

dependents, while the latter sometimes mark recipients, that is, prospective possessors (see Næss, Chapter 38, Lamiroy and Delbecque 1998 and Lander 2004 for discussion).

Somewhat less widespread is the use of the same case for the possessor and the source. One example of a language where the ablative case can code the possessor is Jaru, despite the fact that this language does possess a genitive case:

(25) Jaru (Tsunoda 1981a: 194)
ɲunuŋjɲ-d,u jambi-gu guɲar-u ŋa-ji bajan-i
you(SG).ABL-ERG big-ERG dog-ERG Decl-1SG.ACC bite-PAST
'Your big dog bit me.'

Notably, partial genitive/ablative syncretism is also reconstructed for Proto-Indo-European (Beekes 1995).

Even less common, at least to my knowledge, is the genitive/locative pattern of polysemy. Languages exhibiting it include Budukh (Gilles Authier, pers. com.), Sinhala (see 39.1.4), Krongo, Jaqaru, and Siuslwan.

39.2.4 Diachronic issues

Heine (1997b; 2002) and Heine and Kuteva (2001) note six basic grammaticalization schemas for possessives, namely (X = possessor, Y = possessum):

(i) 'Y at X' Location Schema
(ii) 'Y from X' Source Schema
(iii) 'Y for/to X' Goal Schema
(iv) 'X with Y' Companion Schema
(v) '(As for) X, X's Y' Topic Schema
(vi) 'Y, X's possession' Property Schema

As we have seen, schemas (i)–(iii) do participate in the formation of genitive-like expressions. This is not surprising because all of the corresponding cases, namely locative, ablative, and dative, can provide characterization of not only events but also individuals. Therefore the semantic bleaching of the relations designated by these cases may lead to their use as unmarked adnominal cases. The fourth and the fifth schemas are not found with genitives, while the schema (vi) could imaginably provide a source for a construction where a genitive form appears as a self-standing phrase (as in 18), although I am not aware of any apparent developments of this kind.

The origin of the syncretism of marking of core arguments and possessors is more obscure. Sometimes this may be an outcome of coincidence, as in Finnish, where accusative/genitive syncretism has its origin in a particular phonological change. In other languages such syncretism may have resulted from the development of nominal forms into verbal ones, as in Philippine languages, where

an adnominal dependent was arguably reinterpreted as a dependent of the verb (Starosta et al. 1982). All in all, the usual direction of evolution is seemingly from genitive to core cases rather than vice versa (except for languages with oblique cases functioning as genitives).

Various theories attempted to assimilate possessives to other constructions, such as locative (e.g. most papers in Baron et al. 2001) or subject–predicate (Abney 1987) ones. However, many of the synchronic facts these theories take in support are essentially diachronic (cf. Heine 2002). The possessive should therefore be regarded as a phenomenon of its own, and genitive as a phenomenon of its own inside possessives.

Acknowledgements

The author is grateful to Cindy Kilpatrick, Andrej Malchukov, Barbara Partee, and an anonymous reviewer for their comments on drafts of the paper. This work was supported in part by the NSF under grants No. BCS-9905748 and No. BCS-0418311 to Barbara Partee and Vladimir Borschev and the RGNF grant No. 06-04-00194a.

CHAPTER 40

VARIETIES OF INSTRUMENTAL

HEIKO NARROG

40.1 Formal varieties of instrumentals

40.1.1 Core meaning of instrumental case

Instrumental case typically marks the semantic role of an instrument. Palancar (2002: 32) defines this role as 'the role played by the object the Agent manipulates to achieve a change of state on the Patient'. We can take this as the definition of a typical instrument. The following example from Manipuri shows instrumental case marking such a typical instrument (cf. Chelliah 1997: 125; Bhat and Ningomba 1997: 104–6):

(1) *məhak-nə thaŋ-nə u kəki*
he-NOM knife-INS tree cut
'He cut the tree with a knife'

However, the occurrence of instruments does not depend on the existence of a patient, as its use in the motion event in (2) shows:

(2) *əy-nə bəjar-də-gi bas-nə laki*
I-NOM market-LOC-GEN bus-INS came
'I came from the market by bus'

Lehmann and Shin (2005: 20) accordingly distinguish tools (as in (1)) and vehicles (as in (2)) as the main instrumental functions. The former are the core instruments,

978–0–19–920647–6 41-Spencer-ch40 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 594 of 600 May 16, 2008 19:0

while the latter are treated in some languages as locations. A tool is most typically a body part, particularly the hands, but can also be an artefact, such as a hammer. Material and manner, although often expressed by instrumental markers, should be considered as distinct roles. In some grammatical descriptions the term 'means' can be found, either as a synonym for 'instrument', or as a more abstract category, also including, for example, agents.

40.1.2 Formal varieties of the instrumental

According to Lehmann and Shin (2005: 33–4), seven strategies for the coding of concomitant functions, including instrumental functions, can be distinguished. These are (1) concomitant predication (i.e. converbs, gerunds, coverbs; e.g. Mandarin *yòng* 'use'), (2) adpositional marking (e.g. English *with*, *by*), (3) case marking (e.g. Russian instrumental case), (4) verb derivation (see 40.1.5), (5) incorporation (incorporation of nouns for 'hand' or 'foot' in Yucatec Maya; Lehmann and Shin 2005: 71f.), (6) conversion (e.g. from a noun denoting a potential instrument to a verb, as in English *iron* or *hammer*), and (7) lexical fusion (e.g. *kick* in English, which implies the use of a foot as a tool). Cross-linguistically, all of these strategies are also found for the marking of instrumental functions. Individual languages may have several of them, and may differ with respect to which of the strategies they prefer.

Although this article is part of a handbook of 'case', it is centrally concerned not only with case marking, but also with adpositional marking. For one thing, marking through adpositions is more common than case-marking, as the instrumental is typically not a structurally or syntactically definable case but a semantic case that marks peripheral participants and adjuncts. In this study, I use a sample of 200 languages, which is also used in Narrog and Ito (2007).[1] Among these 200 languages, 82 languages had adpositions with instrumental functions, 52 languages had case marking with instrumental function, and 5 languages had both. Case marking clearly outnumbers adpositional marking only in Australia, New Guinea (Indo-Pacific languages) and non-Indo-European Eurasian languages. Even more crucially, the borderline between adpositions and case markers is fluent and may depend on the theoretical position (Spencer, Chapter 12). This is particularly the case when in a language the same class of morphemes, especially suffixes, marks both core arguments and peripheral participants and adjuncts. For example, Japanese has the essive/instrumental particle (suffix) *-de* which has essentially the same morphosyntactic properties as nominative, accusative, or dative case particles. However, because it typically marks adjuncts, it is viewed as a postposition in syntactic

[1] The sample is based on DV (diversity value) sampling as developed by Rijkhoff and Bakker (1998) and referring to the language classification of Ruhlen (1987).

theories that define case structurally, and as a case marker in other theories. A clear distinction between case suffix and an adposition is only possible in languages which have both case and adpositions with clearly different morphosyntactic properties, such as Russian. Avoiding such problems of classification, many grammatical descriptions do not commit themselves to either label but simply speak of instrumental 'affixes', 'relators', or 'markers' (see also Haspelmath, Chapter 33).

Frequently, a language has more than just one means of marking the instrumental. For example, three of the seven strategies for instrumental marking introduced above were illustrated with English examples. English furthermore has at least two instrumental prepositions, *with* and *by*, which share instrumental functions (cf. Quirk et al. 1985: 699f.; section 40.2.1 below). Instrumental functions may also be shared by two or more case markers, such as the Adessive and Instructive in Finnish (cf. Karlsson 1999: 115–27). On the other hand, it is common for a single marker to have two or more phonologically or morphophonologically conditioned variants. Thus, the Tzutujil (Mayan) preposition *ch* takes the form *cha* before postvelars, and the optional form *chi* before non-postvelar consonants (Dayley 1985: 229). The Waiwai (Kariban) postposition *ya/-way* has the form *ya* without prefix and the form *-way* with prefix (Hawkins 1998: 106). A distinction based on whether the instrumental marks a noun or pronoun can be found in Tariana (Maipurean), where nouns are marked with the instrumental–comitative case clitic *-ne* and personal pronouns with the suffix *-ine* (Aikhenvald 2003: 150). Some languages have instrumental inflection for nouns but not for pronouns (e.g. Yagaria (Trans-New Guinean); Renck 1975: 16f), while the reverse case is not known to me.[2]

40.1.3 Restrictions on the use of the Instrumental

As instrumentally-marked phrases are usually adjuncts, their occurrence is typically not subject to syntactic constraints. However, there are semantic-pragmatic constraints on what kind of nouns can take the role of an instrument. For pragmatic reasons, instruments are typically inanimate. This is the distinctive feature of instruments as opposed to companions and agents (e.g. Nilsen 1973: 112). Thus, if an instrumental marker can mark an animate noun, this is normally a comitative or agentive use. Vehicles are an exception, since animals can be used for transport (e.g. *cross Cuba by horse*). Schlesinger (1979) showed in an experimental study that the cut-off point between comitative marking (animate) and instrumental marking (inanimate) may vary from language to language and is often found in the animal domain.

[2] An exhaustive bibliography on instrumental marking up to the year 2000 can be found in Schwarz et al. (2001).

40.1.4 Distribution of Instrumental: Instrumental on non-typical hosts

Instrumental markers usually govern nouns and noun phrases. In many languages, however, they can extend their domain to govern nominalized verb phrases or clauses as well, resulting in adverbial clauses which have mostly causal or temporal meaning. For example, the Wintu instrumental case suffix *-r* also functions as a subordinating suffix signalling causality or temporal anteriority (Pitkin 1985: 514). In Kwaza, the instrumental case marker *-ko* may suffix to a verb marked by the nominalizer *-ñai* to form causal adverbial clauses (van der Voort 2004: 514). Similarly, the Basque instrumental case *-z* on verb participles may form a temporal, circumstantial, or causal phrase (Lafitte 1962[1978]: 228f.). Instrumental markers on adverbs are less common. In German, instrumental prepositions *durch* and *mit* can combine with the adverbs *da-* and *wo-* to form so-called pronominal adverbs (e.g. *dadurch* 'thereby'), which substitute inanimate noun phrases (Helbig and Buscha 1996: 264–6).

40.1.5 Alternative strategies

Adpositions and case marking form the grammaticalized core of instrumental marking, and because their distinction may depend on theory, they have been treated equally as the subject of this article. However, the semantic role of instrument is also relatively frequently marked by converbs. Converbs are found in at least fifteen languages in the 200-language sample used for this study, most prominently in South East Asia and the Pacific. Commonly, they have meanings such as 'use', 'take', or 'have'. In Tetun Fehan (Austronesian), for example, the prepositional verb *hodi* 'bring, use' features in instrumental phrases such as *k-aré* (1SG-see) *k-odi* (1SG-use) *matan* (eye) 'I see I use eye' (van Klinken 1999: 274). A few languages have instrumental relational nouns, for example the Mam (Mayan) relational noun *-u7n* for instruments (England 1983: 184). Converbs and relational nouns can be considered as stages of grammaticalization in the development of adpositions and case markers. Occasionally, an instrumental marker is labelled as a clitic, for example the Tariana instrumental–comitative case clitic *-ne* (Aikhenvald 2003: 150).

A different strategy is the marking of instrumental function on the verb. One can distinguish two cases:

(1) Applicative affixes: Verbs are derived with instrumental affixes to add instrumental objects external to the verb. In Koasati (Muskogean), the instrumental *s-/st-* derives a transitive verb with an instrumental object from intransitive verbs. For example, the transitive verb *stiltóhnon* 'use' (lit. 'work with') is derived from the intransitive verb *iltóhnon* 'work' (Kimball 1991: 141).

(2) The item affixed to the verb denotes the instrument itself, and usually does not add a participant external to the verb, e.g. Tümpisa Shoshone *mokose* 'smash (with hand)' (*ma-* 'hand', *-kose* 'smash') (Dayley 1989: 95). Typical affixed instruments are body parts such as hand and mouth, followed by object classifiers and natural forces (cf. Palancar 1999).

There are in-between cases, in which a verb is derived with a non-applicative instrumental affix, and still takes an additional external instrumental object (cf. Comanche examples in Charney 1993: 86; explanation in Palancar 1999: 154). The applicative strategy can be found in many language areas (cf. Haspelmath and Müller-Bardey 2004: 1135), while the affixation of the instrument itself is particularly frequent in North American languages (cf. Mithun 1999: 118f.).

40.1.6 Use of case in complex (adpositional, etc.) structures

In languages that have both instrumental case and instrumental adpositions or verb affixes which govern case-marked nouns, the two means of instrumental marking can co-occur. In Russian, the instrumental preposition *s* 'with' governs instrumental case. Besides, the Russian instrumental is also required by a number of other, primarily locative, prepositions (cf. Pulkina and Zakhava-Nekrasova 1997: 104–9). Similarly, Mithun reports cases from Central Pomo (Hokan) in which verbs with instrumental prefixes take nouns with instrumental case marking (Mithun 1999: 120f). According to Gvozdanović (1997), preposition + instrumental case (or other case) is used now in many Slavic languages in various functions where originally the instrumental case was used alone. Instrumental case also participates in case stacking in languages where this phenomenon occurs (cf. Plank 1995a).

40.2 Functional varieties and polysemy patterns

40.2.1 Functional varieties of Instrumental

Instrumental case markers and adpositions usually cover a wide functional domain, and are rarely specialized on the expression of a single function or semantic role. In languages with more than one instrumental marker, the markers may expose subtle semantic and pragmatic differences of use and thus share instrumental functions. For example, Helbig and Buscha (1996: 413–43) list seven prepositions with an instrument function in German. Three of them can be used for vehicles; six of

them with a tool; one of them implies a positive result of the action performed with the instrument; and one can only be used with abstract notions. Each of them, however, is polysemous, and has also other (non-instrumental) functions. As a rule, the more grammaticalized a marker in this area is, the more polyfunctional it is, and the less grammaticalized it is, the higher is the possibility that it is dedicated to a specific grammatical function. Therefore, functional specialization appears to be more common with converbs. For example, in Japanese, 'by X' (X=vehicle) is expressed with the verb phrase X-*ni* (LOC) *notte* ('mount'), lit. 'mounting on X' (cf. Lehmann and Shin 2005: 53), which is not used to express any other grammatical function.

40.2.2 Polysemy patterns

Instrumental markers are typically polysemous. The polysemy in the area of comitative–instrumental has been explored in considerable detail by Stolz and research associates (e.g. Stolz 1996b, 2001a). Based on the 200-language sample used here, markers which have an instrumental function most commonly also mark one or more of the following functions (the numbers show the proportion of markers which additionally have these function among all markers with instrumental function):

(a) Polysemy with syntactic functions: agentive (ergative, agent, passive agent) 0.25, general (direct/indirect/oblique) object 0.18
(b) polysemy with non-locative semantic functions: comitative 0.37, cause/reason 0.29, manner 0.25, material 0.18, ablative 0.14
(c) polysemy with locative functions: location 0.27, point of time 0.19, direction 0.15.

The numbers show that polysemy with other non-locative semantic functions, particularly the comitative, is most common. With respect to core syntactic case functions, polysemy with agentive function is most frequent. The prominent polysemies with comitative and with agentives have received ample attention in the research literature (e.g. Stolz 1996b, 2001a, 2001b; Luraghi 2001; Palancar 2001, 2002).

Some of the polysemies exhibit salient areal and genetic biases. The polysemy between instrumental and comitative appears to be most common in Nilo-Saharan (0.69) and Indo-European (0.58) languages, and in Africa (0.55) and Eurasia (0.53) in general. Polysemy of instrumental with ergative is primarily found in Australia (0.40). Afro-Asiatic (0.54) and Nilo-Saharan (0.46) languages show a strong affinity between instrumental and locative. Polysemy between instrumental and cause/reason is found most often in Indo-European (0.50) and Pama-Nyungan (0.42) languages, and polysemy between instrumental and manner in Eurasia (0.53) and Africa (0.47).

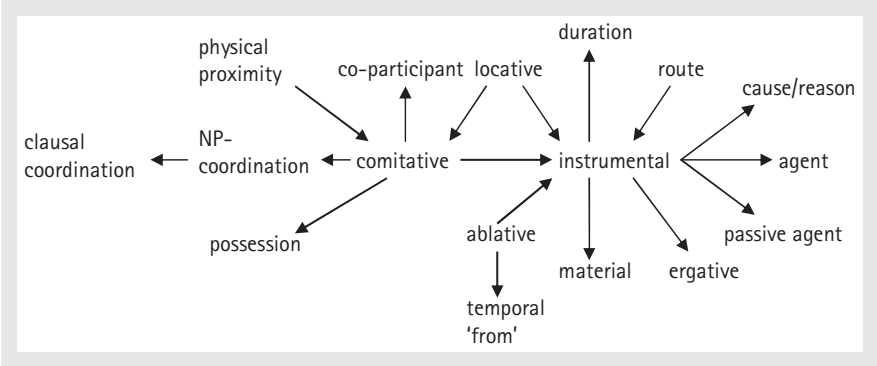


Figure 40.1. Semantic map of the instrumental domain

The data above refer to polysemies which are based on a direct connection between instrumental function and the other function. Instrumental function can also be related to other functions indirectly, via the comitative, or the ablative for example. If these wider connections are taken into account, a semantic map of an area centring on instrumental function can be drawn as in Figure 40.1 (see Narrog and Ito 2007 for more detail). The arrows indicate directionality of functional extension (see section 40.3 below).

The map only contains functions which can be located and connected reliably on the basis of the author's database. In this sense, it contains a minimal number of functions and connections, centring on the instrumental. It is likely that even between the functions shown here more connections exist.[3]

40.3 Diachronic issues

The arrows in Figure 40.1 already indicate the directionality in the extension from one function to another. This information on the directionality was taken from literature on the instrumental area (mentioned above), from hints in grammars on individual languages, and from literature on grammaticalization (Heine and Kuteva 2002). With respect to the most central functions, it is widely agreed that instrumental function, which is essentially inanimate, is derived from comitative function and not vice versa (e.g. Heine et al. 1991: 166; Luraghi 2001, Stolz 2001b). The directionality between instrumental function and the agentive functions (ergative,

[3] Malchukov (p.c.), for example, suggests an additional connection between route and duration, and Stolz (2001b: 340) a second type of possession, which is connected to the instrumental instead of the comitative, via locative function. Both appear to be plausible, but cannot be deduced from the data centred on instrumental function which are available to the author.

agent, passive agent) is more controversial. Palancar (2001), based on a sample of markers from 137 languages, sees a unidirectional development from instrumental to agentive functions (Palancar, however, posits an obligatory intermediate stage of causation), but Luraghi (2001) claims bidirectionality. However, while she attests the directionality from instrumental to agentive functions historically in Indo-European, the case for the opposite directionality (Luraghi 2001: 399) is not based on concrete evidence. Research on the cross-linguistic genesis of ergative systems also supports the directionality from Instrumental to Agentive (Ergative), not necessarily with intervening cause function (cf. Garrett 1990). The development from instrumental to agentive marking is interesting as it means an increase in animacy and is a potential counterexample to commonly posited directionalities in grammaticalization (cf. Heine et al. 1991: 159f.). Furthermore, Palancar (2002: 121–31) suggests bidirectionality between instrument and cause.

Heine et al. (1991: 167) locate case markers and adpositions on a cline between lexical items (nouns, verbs) and zero, that is, loss of case marking. There are various indications in descriptive grammars that this cline applies to instrumental markers as well. Concerning loss of case, Proto-Indo-European is generally believed to have had instrumental case (cf. Szemerényi 1996: 160), but many Indo-European languages have lost it, in some cases while still retaining core cases.

40.4 Lexical/idiosyncratic uses of Instrumental

Like all other grammatical markers, instrumental markers as well can lose morphological productivity and become obsolete, surviving only in specific idioms or lexical items. This appears to be the case with the Finnish instructive, which 'occurs almost exclusively in a few fixed plural expressions' (Karlsson 1999: 127). The instrumental prefixes found in North American languages seem to be particularly prone to develop lexical, idiosyncratic meanings. Palancar presents the lexicalization of prefix–verb combinations as a step in the natural life cycle of such constructions from lexical to grammatical, finally ending up in morphologically intransparent lexicalizations (Palancar 1999: 164).

CHAPTER 41

VARIETIES OF COMITATIVE

THOMAS STOLZ
CORNELIA STROH
AINA URDZE

41.1 Formal varieties of Comitative

In what follows, our observations are based on a world-wide convenience sample of 320 languages (including regional varieties) whose genetic, areal, and typological composition is spelled out in Stolz (1996b) and Stolz, Stroh, and Urdze (2005; forthcoming). The sources from which we draw our examples are mainly the usual descriptive materials with additional data from questionnaires and text analyses. Wherever a fixed norm exists, our examples reflect standard usage of a given language. The discussion of variation and change is reserved to section 41.3. Except otherwise stated, we only take account of grammaticalized dedicated markers of Comitative. Among previous work on Comitative and related issues, the contributions by Seiler (1974), Heine et al. (1991) and Lehmann and Shin (2005) have influenced us most. A specialized bibliography of Comitative (up to the year 2000) is provided by Schwarz et al. (2001). Owing to the frequent syncretism of Comitative and Instrumental, the reader is advised to also consult Chapter 40 on Instrumental which complements this one in many respects (including the bibliographical one). The fact that Comitative and Instrumental go together well also has a bearing on

the validity of what is said in the subsequent sections: wherever a marker is polysemous such that it expresses not only Comitative but also Instrumental, it might be difficult to decide whether certain properties of this marker can be ascribed to the meaning component Comitative or to the one of Instrumental or to the specific combination of both.

41.1.1 Core meaning of Comitative

The core meaning of Comitative is Accompaniment, meaning: a proper Comitative marker is used normally to encode the relation between two participants in an event such that one of them is the accompanee and the other the companion (Stolz 2003: 214–18; Stolz et al. forthcoming: Part A 4.1.1). Accompanee and companion are prototypically human beings. They are involved in the same situation. Their involvement, however, is depicted as asymmetrical as the accompanee is given prominence whereas the companion is somewhat marginalized. This asymmetry is often reflected by grammar because morphosyntactically, the constructions mostly are oriented towards the accompanee. In the Italian sentence *[il professore]*$_{accompanee}$ *entra nell'aula [**con**]*$_{relator}$ *[i suoi studenti]*$_{companion}$ 'the professor enters the lecture-hall (together) with his students' the finite verb agrees with the subject-NP *il professore* 'the professor' in person and number, viz. third person singular. In this situation, this subject-NP represents the accompanee whose companion is the group of students represented by the complement NP *i suoi studenti* 'his students' of the optional PP headed by the Comitative marker *con* 'with'. Thus, only the accompanee-NP has argument status and triggers agreement whereas the companion-NP belongs to an optional adjunct without morphological repercussions.

41.1.2 Formal varieties of Comitative

The inventory of formal means expressing Comitative comprises the following major strategies: (a) affixation (both on [pro]nouns and verbs), (b) adpositional constructions, (c) adverbial constructions, and (d) serial-verb constructions. For the present purpose, we treat (a)–(d) as instances of case relators. In our sample, bound morphology encodes Comitative in 50 per cent of those cases which do not display Comitative–Instrumental syncretism (Stolz et al. forthcoming: Part B, section 10.4). As to affixation, suffixation expectedly predominates (cf. Hungarian *-val/-vel* as in *család**dal*** 'with family'; Lee 1990: 128–9). However, prefixation (cf. Totonac *ta'-* as in ***ta'-****minlh kiamigo* 'He came with my friend'; Levy 1990: 120) and circumfixation (cf. Chukchi *ge-…-e/ga-…-ma* as in ***ge****milger****e****/****ga****melgar****ma***

'with the rifle'; Kämpfe and Volodin 1995: 53–4) are also attested more than just once. Similarly, the adpositional strategies include prepositions (cf. French *avec* as in ***avec*** *sa mère* 'with his/her mother'), postpositions (cf. Wayãpi *lɛwɛ* 'with' as in *yãyã* ***lɛwɛ*** 'with (my older) sister'; Grenand 1980: 62), and circumpositions (cf. Bambara *ní…yé* as in ***ní*** *Músa* ***yé*** 'with Músa'; Kastenholz 1989: 161). There are also instances of combined adpositional and affixal strategies when a given adposition displays differential case government as Icelandic *með* 'with' which requires either the accusative or the dative on its complement (cf. *með þér* 'with you [dative]' vs. *með þig* 'with you [accusative]' where the companion is taking part in the event as an actor only in the former whereas the accusative in the latter PP indicates that the participant is an undergoer; Stolz et al. forthcoming: Part C, section 12.3.2). Adverbial constructions sometimes border on adpositional ones as the morphosyntactic behaviour of the adverb oscillates (cf. Latvian *līdz(i)* 'along' as in *es eju [jums]* ***līdzi*** 'I come along [with you]' where the adverb may also be responsible for the dative case on the pronoun *jums* 'you [dative plural]'; Stolz et al. forthcoming: Part C, section 12.2.2). In serial-verb constructions, Comitative is usually expressed by a co-verb whose basic meaning corresponds to English *to follow* (cf. Chinese *gēn* 'to follow' as in *wǒ* ***gēn*** *tā shuohuà* 'I am conversing with him'; Bisang 1992: 182). Verbal morphology encodes Comitative mostly in languages of the head-marking type. It is exceptional to find NP-agreement involving the Comitative (cf. Sumerian *-da* as case marker on nouns and agreement marker on verbs; Balke 2006). Compared to so-called grammatical case categories, Comitative relatively rarely forms part of a proper inflectional paradigm (thus acquiring the status of a full-blown morphological case as Basque *-ekin* as in *gizonar****ekin*** 'with the man' which contrasts with other case-forms such as ergative *gizonak*, etc.). In many Australian languages, for instance, Comitative is considered a derivational category which does not have a paradigmatic relation to proper inflectional cases in the same language (cf. Yindiny *-ji-* as in *waguja-ŋgu wagal-****ji****-ŋgu dagul-****ji****-ŋgu* 'the man with three wives' where the ergative marker *-ŋgu* is suffixed to the base which already contains the marker of Comitative *-ji-*; Dixon 1980: 325).

41.1.3 Restrictions on the use of Comitative

According to Croft (1991), Comitative is restricted to the role of co-participant of the subject of a clause. However, the empirical data suggest that this restriction is too strong as accompanee–companion constellations are also attested frequently for other fundamental relations and macro-roles, for that matter (cf. German *Erich fotografiert //[die Mutter]*$_{accompanee}$*[****mit****]*$_{relator}$ *[dem Kind]*$_{companion}$*//*$_{patient}$ 'Eric is taking a photo of the mother and/with her child' where the relevant post-verbal constituents form the direct object of the verb and are assigned the patient role;

Stolz et al. forthcoming: Part A, section 4.1.1). In point of fact, animacy has a say in the profile of Comitative. The restrictions very often yield a multiplicity of distinct Comitatives in lieu of simply blocking the category (cf. also section 41.2.1 below). In a variety of languages, both participants are obligatorily either human or animate, at least (cf. Kalmyk *-la/-lä* which presupposes two human participants; Benzing 1985: 68). Elsewhere, the use of Comitative requires accompanee and companion to display identical animacy specifications (cf. Guaraní *ndive* is restricted to animate participants whereas the postposition *reheve* occurs in contexts of reduced animacy; Gauto Bejarano 1990: 114–15). A third group of languages restricts high animacy to the accompanee whereas in group four, no animacy-based restrictions apply at all (cf. Estonian *-ga* in *habemaga mees* 'bearded man' and *jäätі moosiga* 'ice-cream with jam'; Hasselblatt 1992: 72).

41.1.4 Distribution of Comitative

Comitative markers are frequently employed to form conjunctions or to mark subordinate clauses. In Maltese, the Comitative preposition *ma*' 'with' enters a paradigm of conjunctions with Instrumental *bi* 'with', and Benefactive *i* 'for' combining with the general subordinator *li* 'that': *malli* 'when; as soon as', *billi* 'because; while', *illi* 'and thus' (Borg and Azzopardi-Alexander 1995). Similarly, the cliticized allomorph of Tamashek *d* 'with' occurs as (part of the) marker of adverbial, purposive, and subjunctive clauses (Heath 2005: 281).

41.1.5 Alternative strategies

Nevis (1988) discusses the unclear morphological status of Comitative in a number of Uralic languages. For several Saami varieties, there are morphotactic differences between the Comitative in the singular and plural with the latter behaving more like a clitic: cf. *ńiit-**n**-ɛś* 'with his daughter' vs. *ńiit-iś-**k'ejm*** 'with his daughters' where the Comitative marker precedes the possessor affix in the singular but follows it in the plural (Nevis 1988: 48–9) and thus diverges from other case affixes of the same language. In older stages of Sri Lanka Creole Portuguese, the Comitative marker *juntu* had two allomorphs – a postposition and an enclitic (Smith 1978). This situation is also reported for the Turkish postposition *ile* whose allomorphs *-la/-le* closely resemble case affixes as they obey the rules of vowel harmony (Bassarak 1988). Better known examples stem from Latin and some of its Romance offspring: with nouns, Comitative requires an adpositional strategy involving the preposition Latin *cum* (+ $N_{ablative}$), Spanish *con*. However, for pronouns, a distinct inflectional strategy is attested, namely Latin *mecum* 'with me', Spanish *comigo* (Iggesen 2005b: 16).

41.2 Functional varieties and polysemy patterns

If we move away from the prototype of Comitative, described in section 41.1.1, a whole array of potential case categories arises which may pass as functional varieties of Comitative. This is so because neither the companion nor the accompanee is necessarily human or animate. Thus, various combinations of ± human accompanee and ± human companion are conceivable.

41.2.1 Functional varieties of Comitative

In some languages, there is indeed an entire paradigm of distinct Comitative categories corresponding to the different degrees of animacy of the participants (cf. Even with its four distinct Comitatives; Benzing 1955: 64, but see Malchukov 1995). Note that animacy connects easily to the other major parameter determining Comitative, viz. control, because in situations of unequal animacy the participant displaying the higher degree of animacy is normally considered to be in control (cf. Auyana *-kwada* which is used on an NP whose referent exerts control over the event whereas *-yen* has to be used if no control is exerted; Stolz 1997). For a more extended system, see Wilkins (1989) on Arrernte. Minor dichotomies (Stolz et al. forthcoming: Part A, section 4) which may regulate the use of Comitative markers in particular languages are singular/plural, alienable/inalienable, static/dynamic, actor/undergoer, and permanent/temporary. Alongside the more general cover-term Comitative, we also employ special labels for certain segments of the functional domain of non-prototypical Comitative: Confective (transported object) for which Takelma has a special morpheme *-agw-* (Sapir 1922: 138), Ornative (bodily property/clothes) for which French uses *à* 'on, at' instead of *avec* 'with' (Cadiot 1990), Associative (abstract properties) for which Hungarian employs *-stul/-stül* in lieu of *-val/-vel* (Lee 1990). On the semantic map, these categories/functions are situated between the Comitative prototype and possession because they relate to temporary or permanent relations in which the companion is physically or ideally very close to the accompanee and thus almost in his/her possession.

41.2.2 Polysemy patterns of Comitative

The often reiterated universal of Lakoff and Johnson (1980) according to which Instrumental and Comitative are always syncretistic with each other has turned out to be an areal preference of the languages of Europe (Stolz 1996b). In 81 per cent of our sample languages, the relator of Comitative is distinct from that of Instrumental

(cf. Swahili Comitative *na* vs. Instrumental *kwa*) – leaving only a minority of 19 per cent which behave according to the disproved companion metaphor (cf. Faroese Comitative–Instrumental *við*). Even in Europe, Comitative–Instrumental syncretism accounts only for 41.7 per cent of the cases. Of the world's (macro-)phyla, only four attest to a preference for Comitative–Instrumental syncretism, namely Indo-European, Uralo-Yukaghir, Khoisan, and Pidgins and Creoles (Stolz et al. forthcoming: Part B, section 8). Everywhere else, Comitative and Instrumental tend to be kept formally apart. Nevertheless, the relation between Comitative and Instrumental is special as the syncretistic pattern involving both case categories outnumbers all other combinations of either Comitative or Instrumental with different categories: in 32.8 per cent of all cases, our sample languages display formal identity of Comitative and Instrumental (clustering in Europe) whereas the second best pattern is Instrumental–Locative with 24.5 per cent (with a focus in Africa) and in third position, we find Comitative–Coordination (cf. Classical Aztec Comitative *īhuān* 'with, and' vs. Instrumental *īc* 'with') with 15.5 per cent – this pattern has a hotbed in Africa (cf. also Stassen 2000). The only other noteworthy binary combination in which the Comitative partakes is Comitative–Possession with 8 per cent – again with an areal focus in Africa (cf. Swahili Comitative *na* in ***nina*** *kisu* 'I have a knife'; Loogman 1965: 285); all other patterns have marginal frequency (Stolz et al. forthcoming: Part B, section 9). On this basis, it is possible to draw a semantic map which resembles closely the one provided in Chapter 40 (Figure 40.1) – and thus we refrain from reproducing it here. However, this semantic map can be read in the following way: if Comitative syncretizes with a function situated on its right, the syncretistic pattern also involves Instrumental (and vice versa, if Instrumental syncretizes with a function situated on its left, the syncretistic pattern also involves Comitative). These hypotheses are largely corroborated by the empiricial facts – although mostly in the shape of tendencies. If for instance the Comitative marker is used to encode the passive agent, it always also serves the purpose of marking Instrumental (cf. Breton *gant*). The different compatibility of Comitative and Instrumental with agent-oriented categories like ergative and passive agent is described in Stolz (2001a). On the other hand, if the Instrumental marker also expresses co-ordination, it overwhelmingly also has Comitative functions (cf. Fon *kpó [dó]* … *kpó [kpán]*; Höftmann 1993: 138). Thus, we encounter certain types of polysemy of Comitative which show extensions to the right side of the semantic map only in those languages in which Comitative–Instrumental syncretism applies, that is, we are dealing with a minority solution.

41.2.3 Lexical/idiosyncratic uses of Comitative

Occasionally, Comitative markers are used in a way which cannot be predicted from our knowledge of the further usage to which it is put. This may be the effect

of historical merger as in the case of the segmental identity of the Comitative–Instrumental preposition *me* 'with' and the so-called infinitive marker in the Gheg variety of Albanian (Stolz et al. forthcoming: Part C, section 11.1). In Maltese, the Comitative preposition *ma*' and the Instrumental preposition *bi* form a paradigm of temporal expressions: ***ma****l-lejl* 'during the night' vs. ***bi****l-lejl* 'by night' where *ma*' retains some of the prolative meaning associated with other spatial uses of the preposition (Stolz et al. forthcoming: Part C, section 12.1). In a small number of languages, no dedicated grammatical expressions of Comitative seem to exist. An illustrative example is Mapudunguyu – a language which encodes Comitative relations via dual or plural pronouns; similarly, traditional Otomí resorts to dual pronouns to refer to an accompaniment relation of the prototypical kind (Stolz et al. forthcoming: Part D, section 13.2).

41.2.4 Uses in complex structures

Complex structures are relatively common contexts for Comitative markers to occur in. This is straightforward with the Finnish Comitative postposition *kanssa* which governs the genitive on the noun. In Georgian, the postpositional construction is even more complex as it involves the inflectional instrumental *-t* plus the cliticized postposition *gan* and the case-inflected numeral *erti* 'one' in the adverbialis case: *N-t + gan + ert-ad > N-dan ertad*. Moreover, in Finnish, the inflectional Comitative *-ine-* can never occur on its own but always combines with the appropriate possessor suffixes. In numerous languages of Europe, the expression of Comitative is at least binary and involves the expression of Instrumental. Whereas Instrumental is expressed by morphologically simple and primary means (such as case affixes), Comitative often requires at least one additional morpheme (usually an adposition) as in Latin (Instrumental = $N_{ablative}$ vs. Comitative = *cum* $N_{ablative}$), Lithuanian (Instrumental = $N_{instrumental}$ vs. Comitative = *su* $N_{instrumental}$), Armenian (Instrumental = $N_{instrumental}$ vs. Comitative N_{dative} *het*), and many members of the Slavic phylum (Instrumental = $N_{instrumental}$ vs. Comitative *s* / *z* $N_{instrumental}$), etc. This pattern of unequal complexity of constructions recurs so often that attempts have been made at understanding the relation in terms of markedness (Stolz 1998).

41.3 Diachronic issues

In the framework of grammaticalization research, Comitative occupies a very important position because many of the early hypotheses of this branch of linguistics are based on the validity of the companion metaphor (Heine et al. 1991). Moreover,

978–0–19–920647–6 42-Spencer-ch41 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 608 of 608 May 16, 2008 19:0

in various publications, for instance Heine (1997a), Heine postulates a cognitive schema labelled Comitative schema which belongs to the basic inventory of patterns which underlie the constructions of countless grammatical categories in human language(s). The internal logic of the grammaticalization model is such that the schemas can be interpreted diachronically in the sense that a realized structure corresponding to the cognitive schema may develop into the construction for expressing certain functions. Thus, the semantic map alluded to above can be understood as a scenario for the diachronic development Comitative > Instrumental, for instance. This grammaticalization path is amply attested for the Baltic region (Stolz 2001c). However, the diachrony of Comitative also has a circular variant. In the early days of Estonian, Instrumental was expressed by a reflex of the old inflectional instructive whereas Comitative required a postpositional construction (with the postposition *kaes*) which subsequently was morphologized into a new case affix *-ga*. The newly developed Comitative case took over the functional domain of the erstwhile instructive and thus became a Comitative–Instrumental case. In contemporary Estonian, however, an additional free morpheme *koos* 'together' is increasingly used either postpositionally or adverbially or prepositionally in order to distinguish Comitative from Instrumental readings of the polysemous case affix. The more complex construction is exclusively used for Comitative (Stolz et al. forthcoming: Part D, section 13). This diachronic development has parallels in neighbouring Latvian and several other languages not only in the European east. As to the lexico-semantic sources of Comitative markers, wherever the etymology can be ascertained, there is a clear preference for lexemes which cover the following broad meanings: 'company' (cf. Finnish *kanssa* < *kansassa* 'in the company [of]'), 'be one with' (Estonian *ühes* 'in one'), 'body' (cf. Baka *tɛ* 'body'), 'join', 'meet', 'follow', 'accompany', etc. (Givón 1979, Hagège 1993, Heine/Kuteva 2002). All of these concepts have affinities to socio-spatial proximity and thus define Comitative conceptually as the relation of co-presence and co-involvement in a given situation (Stolz et al. forthcoming: Part D, 13.1). This conceptual basis explains why Comitative is prone to acquire functions of coordinating conjunctions and those of possession as these categories both also presuppose co-presence and/or co-involvement of entities.

CHAPTER 42

SPATIAL CASES

DENIS CREISSELS

42.1 Introductory remarks

42.1.1 Definitions

In this chapter, the term *case* is taken in its traditional meaning of inflectional category-system (and the individual categories or values of that system) expressing dependency relations involving NPs.[1]

A spatial relation involves two percepts, a *Figure* (or *Theme*, or *Trajector*) and a *Ground* (or *Location*, or *Landmark*), the Figure being perceived as located or in motion relative to the Ground.[2]

A spatial case is an inflected form of nouns or NPs distinct from the absolute form available for the extra-syntactic function of pure designation, and apt to fulfil one of the following functions without the addition of an adposition:

- non-verbal predicate, or predicative complement of a copula, specifying the location of an entity
- verb satellite specifying the location of an event
- argument of motion verbs specifying the source, path, or destination of the movement.

[1] For a discussion of the various extensions of the term *case* encountered in the literature, see Haspelmath (Chapter 33).

[2] For a general approach to the study of the linguistic expression of spatial relations, see e.g. Jackendoff (1983), Langacker (1987), Jackendoff and Landau (1992), Svorou (1994), Pederson (1995), Pederson et al. (1998), Talmy (2000). Shay and Seibert (2003) provides a collection of papers exploring the variety of the linguistic means of expressing spatial relations in typologically diverse languages.

It may happen that the same case form has some uses corresponding to this definition along with non-spatial uses (see sections 42.4–5).

Some authors have advocated that all cases can be viewed as locative expressions (see in particular Hjelmslev 1935, 1937, Anderson 1971, Cook 1989). Since we are concerned with cases that have concrete spatial meanings (and not with cases having only non-spatial uses derivable from abstract locative meanings) this question will not be further developed here.

42.1.2 Simple and complex spatial cases

Morphologically complex spatial cases may include a formative having an independent existence as a spatial case marker, combined with an *extension*, that is, a formative that has no independent existence (for example, Basque directional *-ra-ntz* 'towards' and terminative *-ra-ino* 'up to' combine allative *-ra* with two extensions that exist only in combination with the allative marker).

It may also happen that none of the formatives involved in the formation of a complex spatial case marker has an independent existence. In Akhvakh, all spatial cases arise from a combination of markers for relative orientation with markers for destination, location, or source, but none of the distinctions underlying this system involves zero marking.

Spatial cases arising from a combination of three formatives are attested. In some Daghestanian languages, cases markers encoding path consist of a first formative for relative orientation, a second formative that by itself carries an ablative meaning, and a third formative, responsible for the meaning of path, that exists only as an extension of the ablative marker. In Tsez, the possible addition of a distal marker is responsible for the existence of spatial cases consisting of four formatives (Comrie 1999).

42.1.3 Problems in separating spatial cases from spatial adpositions

It is not possible to discuss here the criteria according to which descriptive linguists decide to treat simple markers encoding spatial meanings and located at the edge of noun phrases as spatial cases or spatial adpositions. In the case of complex markers, three options are available (combination of two affixes, of an affix and an adposition, or of two adpositions), and the criteria that may guide a choice between these three options cannot be discussed here either:

> Although one can easily separate different layers of case marking in a particular language, as in Hindi for instance, it can be difficult to determine whether a single layer of case marking in a particular language is affixal or adpositional. (Blake 1994: 11)

Consequently, in the search for typological generalizations concerning spatial cases in the narrow sense of this term, one must always keep in mind that there is some degree of arbitrariness in the distinction between cases affixes and adpositions as it is recognized in the descriptions of individual languages. It is nevertheless interesting to take this distinction into account, because spatial relation markers that are clearly affixal lend themselves to some generalizations that do not apply to those that are clearly adpositional and vice versa.

42.1.4 Spatial cases in languages devoid of case contrast between core syntactic terms

Most of the time, noun forms identifiable as spatial cases according to the definition put forward in the introduction are observed in languages in which morphological variations of nouns contribute to marking the contrast between syntactic core terms (*nominative* vs. *accusative*, or *absolutive* vs. *ergative*), but there are exceptions. Tswana has no case contrast between syntactic core terms, but Tswana nouns have a locative form. Nahuatl is another case in point.

42.1.5 Case languages devoid of spatial cases *stricto sensu*

A language may have case inflection without having spatial cases. For example, none of the three cases of Classical Arabic (nominative, accusative, and genitive) can be used by itself in spatial functions.

In case languages devoid of spatial cases proper, the functions typical of spatial cases are expressed by means of adpositional phrases. This is a common situation among the modern Indo-European languages that have not entirely eliminated the Indo-European case inflection. For example, in Russian and other Slavonic languages:

- the ancient locative case has become a *prepositional case*, that can only be used in combination with some prepositions
- the case resulting from the fusion of Indo-European genitive and ablative still assumes the genitive function by itself, but assumes the ablative function in combination with some prepositions only
- similarly, the accusative case, which originally assumed the allative function by itself, occurs in allative function in combination with prepositions only.

Case forms used in such conditions may still contribute to the expression of spatial meanings (in Russian, both *za* + *accusative* and *za* + *instrumental* encode the relative orientation BEHIND, but the choice between accusative and instrumental

case encodes the distinction between destination and location).[3] It may however also happen that the choice of a particular case is simply required by the adposition (in Russian, *ot* 'from' is necessarily followed by a noun phrase in the genitive case).

In the old Indo-European system, most clearly attested in Sanskrit, three case forms of nouns qualified as spatial cases, at least in some of their uses. Latin illustrates the transition towards another type, in which the same case forms can contribute to the expression of spatial meanings within the frame of prepositional phrases only: in Latin, with some nouns, the accusative case assumed an allative function, and the ablative case assumed locative and ablative functions, without the help of any preposition,[4] but for most nouns, the expression of spatial meanings necessitated the presence of a preposition.

In languages that undergo such an evolution, it is common that isolated spatial case forms of some nouns, when they are not accompanied by noun dependents, continue to be used by themselves with spatial meanings, becoming thus spatial adverbs.

42.1.6 Spatial cases and semantic classes of nouns

Two semantic classes of nouns frequently have particularities in relation with spatial cases: Geographical names,[5] and nouns referring to humans.

Geographical names often have a 'lighter' spatial marking than most other nouns and tend to be more conservative in evolutions affecting the expression of spatial relations. This is quite obviously the consequence of their predisposition to represent the reference point in a spatial relation, and of the frequency of their use as spatial complements or adjuncts. In Latin, the nouns that maintained spatial uses of prepositionless ablative and prepositionless accusative were mainly town names. In Tswana, names of towns or countries have no locative form, and occur in the absolute form in contexts in which, with very few exceptions, other nouns must take the locative form. In Hungarian, some town names maintain an ancient locative ending *-ett/ött/ott* that has been eliminated from regular noun inflection.[6]

Nouns referring to humans, or more generally to animate beings, may have special forms for spatial cases. In some languages, such nouns are incompatible with spatial cases (for example, in Armenian, they do not have locative case). This follows from a general tendency to express spatial relations with human beings as the reference point in an indirect way, through a genitival construction ('at N's

[3] Abraham 2003 discusses case assignment with respect to the doubly governing prepositions in German, showing the weaknesses of the traditional presentation of this question.

[4] Some of these nouns (but not all) maintained also a locative form distinct from the ablative.

[5] Common nouns characterizable as 'natural locations' (such as *house*, or *village*) often show the same tendencies as geographical names with respect to the expression of spatial relations.

[6] This ancient locative suffix subsists also in the inflection of spatial postpositions – see section 42.3.

place'). The grammaticalization of such constructions may lead to the creation of spatial adpositions specifically used with human nouns (like French *chez* < Latin *casa*), but in some variants of this scenario, the result may be the emergence of spatial cases specifically used with human nouns; one may for example imagine the grammaticalization of a construction in which a spatial case affix is attached to a genitive interpreted as an elliptical variant of the construction *N's (place)*.

Basque cumulates both particularities: contrary to the other cases of Basque, which have the same form for all nouns, spatial cases have special forms (a) for animate nouns, and (b) for geographical names. Interestingly, the spatial case endings of Basque are shorter for geographical names than for ordinary nouns, and longer for animate nouns.

'Endingless locatives' (which in at least some cases can be analysed as bare noun stems in locative function) are well-attested in some ancient Indo-European languages, but the nouns that have this particularity do not seem to constitute a natural semantic class (for Hittite, see Neu 1980), and the very notion of 'endingless locatives' in Indo-European is not entirely clear.

An interesting case of grammaticalization of the relation between semantic classes of nouns and the encoding of spatial meanings is attested by Central Bantu languages, in which affixes that are basically noun class prefixes (and as such, primarily attach to subsets of nouns that have in common an inherent locative meaning) are also used with nouns belonging to other classes with a function similar to that of spatial case markers (Grégoire 1975).

42.1.7 Adverbs and adpositions inflected for spatial cases

In languages in which nouns have a case inflection including a subsystem of spatial cases, spatial adverbs and adpositions commonly have possibilities of variation similar to noun case inflection, but limited to spatial cases. This can be viewed as an indication that, diachronically, such adverbs or postpositions originate from nominal forms inflected for case – see section 42.3.

42.2 Semantic distinctions expressed by spatial cases

42.2.1 Preliminary remark

When analysing the semantic distinctions expressed by spatial cases, cases that assume spatial functions in a productive way along with uses typical of some

non-spatial case must be considered on a par with specialized spatial cases, and in this respect, current terminology may be misleading. For example, Turkish provides a typical illustration of the tripartite distinction *locative/allative/ablative*, but the case assuming the allative function is also used as a dative marker, and is designated simply as 'dative' in many Turkish grammars, whereas in other languages which have basically the same situation, a case very similar to Turkish 'dative' is described as an 'allative' or 'directive' case having also dative functions.

42.2.2 Case systems with a unique spatial case

Some languages have a unique spatial form of nouns used equally, without the help of any adposition, for verb dependents expressing location, source of movement, or destination of movement (Tswana, Nahuatl). Such forms are commonly designated as *locative*.[7]

But there are also languages with a unique spatial case, designated as *locative*, expressing location only. For example, Baltic languages (Latvian, Lithuanian) encode location by means of prepositionless locative, but use prepositions governing non-spatial cases for destination or source.

42.2.3 Types of distinctions structuring the systems of spatial cases

Spatial case systems tend to be structured along two dimensions: the distinction between location, destination, source, and path on the one hand, and reference to particular types of *spatial configurations* (or *relative orientations*) on the other hand. The terms *configuration* and *relative orientation*, taken here as synonymous, refer to distinctions of the type expressed by the choice between *in*, *on*, *at*, *behind*, *under*, etc. in English.

For a formal analysis of the semantics of locative expressions, see Kracht (2002).

42.2.4 Unidimensional spatial case systems

Spatial case systems encoding distinctions between types of configurations only, without any distinction between location, destination, source, and path, do not seem to be attested.

Unidimensional spatial case systems tend to be organized according to a tripartite distinction between location, destination of movement, and source of movement.

[7] Creissels (2006a) examines the strategies used to encode the distinction between location, destination, and source, in languages that do not encode this distinction by means of adpositions or case affixes. See also Wälchli and Zúñiga (2006).

Locative/allative/ablative are the terms most commonly used to describe such systems, at least when the case forms in question are used predominantly in spatial functions.[8] Another possible terminology, mostly used however in descriptions of bidimensional spatial case systems, is *essive/lative/elative*.[9]

In systems of spatial adpositions, the conflation of location with destination is quite common. Among spatial case systems, this conflation is found for example in Kryz, but this is not a common situation, which is somewhat unexpected, given the diachronic relation between case affixes and adpositions.

A few languages have a *locative-ablative* vs. *allative* distinction: Kanuri, Dinka, Old Georgian; in Latin, only some of the nouns that maintained spatial uses of the ablative case had a distinct locative form. Contrary to what we observed in relation to the location–destination conflation, this second type of conflation is equally rare in case systems and in adposition systems.

42.2.5 Bidimensional spatial case systems

Bidimensional spatial case systems combine an indication on relative orientation with a distinction of the type *locative/allative/ablative* (or *essive/lative/elative*).[10] The indication on relative orientation carried by spatial cases is most of the time limited to three basic configurations that can be symbolized as IN, ON, and AT.[11]

Chart 42.1 gives the terms used to label the nine spatial cases of Hungarian analysable as combining the two dimensions *location/destination/source* and *IN/ON/AT*.[12]

More elaborated bidimensional systems of spatial cases are found in Daghestanian languages[13] and in Burushaski. In such systems, the second dimension may

[8] *Allative* and *ablative* are used with a more restricted meaning in the description of some bidimensional systems of spatial cases – see section 42.2.5

[9] Note that *essive* is also used, for example in Hungarian grammars, as a label for non-spatial cases with the meaning 'as', 'in capacity of'.

[10] Recent works on Daghestanian languages, which have particularly elaborated systems of spatial cases, use the terms *localization* for what is called here configuration or relative orientation, and *direction* for the choice between location, destination, source, and path.

[11] This is in accordance with a general principle according to which, 'Where inflectional case and adpositions co-occur in a language, the adpositional system normally exhibits finer distinctions than the inflectional system' (Blake 1994: 11).

[12] In a consistent use of Latin roots, *sublative* should be replaced by *superlative*, since *sub-* refers to UNDER-configuration. The choice of this term by Hungarian grammarians was probably motivated by the desire to avoid homonymy with *superlative* as an inflected form of adjectives.

[13] On Daghestanian case systems in general, see Daniel and Ganenkov, Chapter 46. On Daghestanian spatial case systems, see Comrie (1999), and for a diachronic approach Alekseev (2003). Detailed presentations of Daghestanian systems can be found in Authier (to appear) (Kryz), Charachidzé (1981) (Avar), Haspelmath (1993a) (Lezgian), Kibrik (1996) (Godoberi), Kibrik et al. (2001) (Bagvalal), Kibrik and Testelec (1999) (Tsakhur), Sumbatova and Mutalov (2003) (Dargwa), van den Berg (1995) (Hunzib).

Chart 42.1. Hungarian spatial cases

	IN	ON	AT
location	*-ban/ben* (inessive)	*-(o/e/ö)n* (superessive)	*-nál/nél* (adessive)
destination	*-ba/be* (illative)	*-ra/re* (sublative)	*-hoz/hez/höz* (allative)
source	*-ból/ből* (elative)	*-ról/ről* (delative)	*-tól/től* (ablative)

include a fourth term specifically used to encode path (*translative*, or *perlative*), and the first dimension may include additional distinctions, for example :

- series encoding configurations more commonly encoded by means of adpositions or locational nouns: UNDER, BEHIND, or IN FRONT OF
- distinct series for two varieties of IN-configuration, 'in filled, dense space' (*The stone is in the ground*) vs. 'within an empty or closed space' (*The water is in the glass*)
- distinct series for two varieties of ON-configuration, according to distinctions such as *vertical* vs. *horizontal contact*, or *contact due to gravity* vs. *adherence* (see Ganenkov 2005 on Daghestanian languages, Tiffou 1999: 190 on Burushaski).

Chart 42.2 gives possible labels for the spatial cases of Avar as they are currently presented in descriptions of Avar, with five series of four cases each.[14] The semantic complexity of such spatial case markers is more or less reflected in their form.

In Hungarian it is not possible, in a strictly synchronic analysis, to isolate in each spatial case marker a formative for the distinction *location/destination/source* and another for the distinction *IN/ON/AT*, but the superessive marker is the only one constituting the direct reflex of an ancient case marker, and most of the others (in

[14] This inventory calls for the following two remarks, which unfortunately cannot be sufficiently developed here:

(a) The first series, traditionally described as encoding ON-configuration, tends to take the status of series unmarked for configuration, the postposition *t'ad* being used to encode specifically ON-configuration.

(b) The basic meaning of the two series labelled here IN_1 and IN_2 is currently described as involving the distinction *in dense space* vs. *in empty space*. <*b*> in the first two forms of the fifth series is a gender–number marker agreeing with the absolutive term of the construction in which these forms fulfil the function of locative adjuncts or allative complements, conventionally given in the form it takes for 'singular, non-human'. This last series has no formative of its own, but its formation involves a special stem, whereas the suffixes of the other series attach to the same 'oblique stem' as non-spatial case markers (for example, the oblique stem of *ruq*' 'house' is *ruq'aɬ-*, but the stem of this word for the suffixes of the fifth series of spatial cases is *roq'o-*). In addition to that, this series exists only with a limited set of nouns, which casts some doubts on the decision to analyse these forms as inflected forms of nouns rather than locative adverbs derived from nouns. Diachronically, it seems reasonable to assume that this fifth series represents a layer of spatial case marking anterior to the emergence of the bidimensional system of spatial cases. Additional evidence supporting this hypothesis is the similarity between this series and the case inflection of adverbs/postpositions such as *hani*<*b*> 'here', *ki*<*b*> 'where?', *horL'o*<*b*> 'in the middle (of)', *χadu*<*b*> 'behind', *žani*<*b*> 'inside' (among which some at least may be ancient locational nouns subsisting only in spatial case forms).

Chart 42.2. Avar spatial cases

	ON	AT	IN_1	UNDER	IN_2
location	*-da* (superessive)	*-q* (apudessive)	*-ɬ* (intereressive)	*-L'* (subessive)	*-<b>* (inessive)
destination	*-de* (superlative)	*-qe* (apudlative)	*-ɬe* (interlative)	*-L'e* (sublative)	*-<b>-e* (illative)
source	*-dasa* (superelative)	*-qa* (apudelative)	*-ɬa* (interelative)	*-L'a* (subelative)	*-sa* (inelative)
path	*-dasan* (supertranslative)	*-qan* (apudtranslative)	*-ɬan* (intertranslative)	*-L'an* (subtranslative)	*-san* (intranslative)

particular, those encoding IN-configuration) still have forms suggesting that they originally combined two distinct formatives.

In Avar, three series of spatial cases quite transparently result from the concatenation of a configuration marker and a marker encoding the *location/destination/source/path* distinction, but the segmentation of the remaining two series is more problematic. In such cases, it can be assumed that evolutions have blurred the original composition of complex markers, or that the system conflates distinct diachronic layers of case marking (see in particular note 14 on the fifth series of spatial cases of Avar).

A plausible source of such systems is the morphologization of constructions involving a case-marked locational noun or adposition expressing configuration – see section 42.3.

In bidimensional systems in which spatial case markers can be more or less straightforwardly segmented into two formatives, it may happen that either:

- the absence of the second formative encodes location, contrasting with non-void formatives encoding source and destination (Avar, Burushaski), or
- the configuration marker is optional, and its absence indicates that the relative orientation is not specified (such a system is reconstructed for Proto-Finnish – see section 42.6).

A frequent problem in describing bidimensional spatial case systems is that one of the series may tend to lose its specific meaning, and to function as the default choice when the specification of a particular configuration is not important from a communicative point of view. For example, Akhvakh has five series of spatial cases; three of them quite obviously encode specific configurations (IN, AT, and UNDER), but the other two have uses that lead to the conclusion that one of them (characterized by a formative *-q-*) has lexicalized (in the sense that its productivity is very limited, and the nouns that have this series of spatial cases must be listed in the lexicon), whereas the other (characterized by a formative *-g-*) is the default

choice to encode spatial relations without referring to a particular configuration. A similar problem has been mentioned for Avar in note 14.

In the absence of a series used by default when no particular configuration is intended, the choice of a particular series may involve lexical properties of nouns. Kracht (2005: 150) rightly observes that Hungarian has *a mezőn* 'in the meadow' (superessive) but *a kertben* 'in the garden' (inessive), and that a *hajón* 'on the ship' (superessive) 'is the neutral way of saying that you are on the ship, regardless of whether you are in a cabin or on deck'.

42.2.6 Tridimensional spatial case systems

Tridimensional spatial case systems are exceptional, but Tsez attests the possibility of combining the two dimensions commonly found in Daghestanian systems with a binary distinction ±*distal*.

42.2.7 The expression of path

In addition to the common tripartite distinction *location/destination/source*, a few languages (e.g. Yup'ik) have a fourth unanalysable marker specifically used to encode path (*translative*, or *perlative*), but the following situations are much more common:

- Ablative markers are often used to encode, not only source, but also path (Basque, Turkish, Armenian, Lezgian).
- Path may also be encoded by a complex case marker consisting of an ablative marker plus an extension in the sense of section 42.1.2 (Avar and other Daghestanian languages).
- Path may be encoded by means of the combination of a spatial case and an adposition (Hungarian).
- Instrumental case markers or adpositions often have the expression of path as one of their possible uses, even in languages that have a distinction between an instrumental case/adposition and an ablative case (in Azerbaidjani, at least with some nouns, the ablative suffix *-dAn* and the comitative–instrumental postposition *ile* can equally encode the role of path).

42.2.8 Other uncommon types of spatial cases

Páez has a unique allative case *-na* and a unique ablative case *-hu*, but four locative cases expressing an obligatory distinction between four types of posture: standing (*-te*), lying (*-ka*), hanging (*-khe*), and leaning (*-su*).

In languages that have spatial cases, *terminative* ('up to') is commonly encoded by means of an adposition governing the allative case, or a complex case marker resulting from the addition of an extension to an allative marker (Basque *-ra-ino*, Azerbaijani *-(y)A-cAn*), but Hungarian has an unanalysable terminative marker (*-ig*). A terminative case is also mentioned in Estonian grammars, but it is a complex marker based on the genitive case.

Directional ('towards') is encoded in Basque by means of an extension of the allative marker.

42.2.9 Common types of spatial cases in uncommon types of organization

From a strictly synchronic point of view, Georgian can be analysed as having an inventory of seven spatial cases.[15] All of them are complex markers ('cas secondaires' in Vogt's terminology), but their meaning is entirely determined by the second formative (which is historically a cliticized postposition), and their first formative is a non-spatial case initially governed by the postposition from which the second formative originates:

- superessive-superlative: dative + *-ze*
- inessive-illative: dative + *-ši*
- adessive-allative: dative + *-tan*
- ablative: genitive + *-gan*
- elative: instrumental + *-dan*
- directional: genitive + *-k'en*
- terminative: adverbial + *-mde*

42.3 Spatial case forms of locational nouns, spatial adpositions, and spatial adverbs

In languages that have unidimensional systems of spatial cases, relative orientation is often transparently expressed by means of locational nouns inflected for spatial cases and taking as their complement a noun phrase (often in the genitive case), as in Basque (Chart 42.3).

[15] See Vogt 1971; most Georgian grammars do not consider these complex markers as cases, but are not explicit on the nature of their second formative.

Chart 42.3. Basque locational phrases

	gain 'top'	*azpi* 'bottom'	*ondo* 'side'	etc.
location	*gain-ean*	*azpi-an*	*ondo-an*	
destination	*gain-era*	*azpi-ra*	*ondo-ra*	
source	*gain-etik*	*azpi-tik*	*ondo-tik*	

The grammaticalization of such situations may lead:

- in a first stage, to systems of spatial postpositions inflected for spatial cases;
- in a second stage, to bidimensional systems of spatial cases.

This is precisely what occurred in Hungarian. This language has a large inventory of postpositions with a case inflection consisting of three forms encoding the distinction *location/destination/source* (Chart 42.4)

Chart 42.4. Hungarian postpositions

	BESIDE	FRONT	BACK	etc.
location	*mell-ett*	*el-ött*	*mög-ött*	
destination	*mell-é*	*el-é*	*mög-é*	
source	*mell-ől*	*el-ől*	*mög-ül*	

These postpositions are quite obviously, from a diachronic point of view, spatial case forms of locational nouns. But synchronically, due to relatively recent changes in the case system of Hungarian, they cannot be recognized as such, even when the noun they originate from still exists (for example, *mell-ett/-é/-ől* is clearly cognate with *mell* 'chest', but none of the three forms of this postposition coincides with any of the case forms of the noun *mell*).

The cliticization of some of these postpositions resulted in the two-dimensional spatial case system of Hungarian presented in section 42.2.5.[16]

In a variant of this scenario, nominal forms or phrases including spatial case markers grammaticalize in contexts in which they are not combined with a complement NP, becoming thus spatial adverbs, and the adverbs created in this way may maintain a case inflection limited to spatial cases, and often irregular from

[16] Note that the Hungarian postpositions whose cliticization gave rise to the case system of present-day Hungarian still exist as words in complementary distribution with case suffixes: in Hungarian, with the only exception of the accusative, case suffixes cannot attach to personal pronouns, and this impossibility is compensated by the use of postpositions inflected for person. For example, 'by me' cannot be expressed as **én-nél* (*én* 'I', *-nél* 'adessive'), but only as *nálam*, first person of a postposition cognate with the case suffix *-nál/nél*; this postposition differs from ordinary postpositions in that it exists only in forms having a personal ending – for a detailed analysis, see Creissels (2006b).

a synchronic point of view, due to the fact that, once they have frozen as adverbs, these forms are not affected by subsequent changes affecting regular noun inflection.

For example, Basque has three deictic place adverbs, *hemen/hon-* 'here', *hor* 'there (near hearer)', and *han* 'over there', with the following possibilities of case inflection:

locative	*hemen*	*hor*	*han*
allative	*hona*	*horra*	*hara*
ablative	*hemendik*	*hortik*	*handik*

The etymological connection with the three demonstratives (*hau(r)/hon-*, *hori*, and *hura/ha-*) is clear, but in the present state of the language, these forms cannot be recognized as case forms of demonstratives, or of phrases including demonstratives.

Adverbs inflected for spatial cases, often cognate with inflected postpositions, are very common in Hungarian. For example, the postposition *el-ött/el-é/el-ől* expressing the relative orientation FRONT is cognate with the adverb *elöl/elő-re/elöl-ről* expressing exactly the same meaning. Akhvakh is another typical example of a language with a large inventory of such adverbs.

42.4 Syncretisms between core syntactic cases and spatial cases

42.4.1 Allative–dative syncretism

Not surprisingly, given the prototypical meanings currently ascribed to cases, the allative–dative syncretism is by far the most common syncretism involving a core syntactic case and a spatial case. Even in languages that have a specialized dative marker, it is not rare that allative can replace dative in its most typical function (e.g. Akhvakh), and the shift *allative* → *dative* is admittedly a common source of dative markers.

42.4.2 Allative–accusative syncretism, and unmarked allative

An *allative–accusative* syncretism in languages that have a specialized dative case is attested in ancient Indo-European languages (Sanskrit, Latin).

A possible explanation is a first change by which an allative marker extends its use to dative and subsequently accusative functions (as attested in the history of the Romance preposition *a* < Latin *ad* in Spanish, or in Sardinian), the chain *allative–dative–accusative* being subsequently broken by the creation of a new dative

marker from a different source. However, this hypothesis is entirely speculative, and the following observations on *unmarked allative* suggest rather a direct functional explanation of such situations.

Unmarked allative (i.e. the use of the absolute form of nouns in allative function without the help of any adposition) is observed in several languages that have also direct objects devoid of any mark of their function (for example, in Maale, an Omotic language in which a marked case form of nouns is used for subjects, and the absolute form is used for objects). The situation of such languages is therefore functionally similar to that of languages using the same case-marked form of nouns for direct objects and complements of verbs of movement expressing destination. In some cases (for example, in Armenian), the unmarked allative can be explained by the loss of the *nominative* vs. *accusative* distinction in a system originally characterized by an *allative–accusative* syncretism. But unmarked allative is also found in languages in which there is no evidence that an accusative case ever existed, and other explanations must be considered.

For example, Tswana has a locative form of nouns indistinctly used for nouns in the roles of location, source, and destination, but the verbs of movement assigning to their complement the role of destination can also optionally take complements encoding destination that bear no mark of their role, and are treated in fact as direct objects (in particular, they can pronominalize in the same way as direct objects).

The tendency to treat nouns encoding destination or source of movement as the second core term of transitive constructions is sporadically attested in many languages (as in English: *He entered/left the room*), and seems particularly strong for nouns in allative function, which may explain the situation of languages in which unmarked allative is the rule.[17]

42.4.3 Locative–dative syncretism

Mongolian has a *locative–dative* syncretism, but uses a different marker (traditionally analysed as a postposition governing the absolute form of nouns) for allative. A similar situation existed in Old Hittite.

42.4.4 Spatial cases and genitive

The *ablative–genitive* syncretism is common in adposition systems (French *de*), and therefore would be expected to be found in case systems too, but examples are not easy to find. Such a syncretism is postulated in the history of Greek, but in Classical Greek, the ablative–genitive case was no more a spatial case, since it assumed spatial functions productively in combination with prepositions only.

[17] More generally, on the possibility of treating NPs encoding spatial relations as core syntactic terms, see Dimmendaal (2003).

The same can be said of the *locative–genitive*, *allative–genitive*, and *locative–allative–genitive* syncretisms: such syncretisms are relatively common in adposition systems (for example, Bulgarian *na* illustrates the *locative–allative–genitive* syncretism), but seem to be rare in case systems.

42.4.5 Spatial cases and ergative

Dixon (1994) signals the existence of a *locative–ergative* syncretism in some languages, but does not mention the possibility of an *ablative–ergative* syncretism. This is somewhat unexpected, since *ablative–instrumental* and *instrumental–ergative* are among the most common case syncretisms, and in many languages, obliques representing agents (in particular, in passive constructions) occur in the ablative case. An illustration of the *ablative–ergative* syncretism can however be found in Kryz, where the ergative suffix coincides with the second formative of spatial cases encoding source.

42.5 Other uses of spatial cases

42.5.1 Oblique arguments

Arguments of individual verbs may occur in case forms that have otherwise the status of spatial cases. Such case assignments can often be viewed as resulting from a metaphorical extension of the spatial meaning of the case in question. For example, verbs with the meaning 'be afraid (of)' commonly assign ablative case to their complement.

42.5.2 Spatial cases in constructions involving valency changes

Ablative is used in some languages for NPs representing the demoted subject in passive constructions (Greenlandic). This use is clearly related to the ablative of cause – see 42.5.3.

Yup'ik uses the ablative case for demoted objects in antipassive and applicative constructions, and the allative case for demoted transitive subjects in constructions involving the introduction of an additional participant in subject role.

Spatial cases may be involved in non-volitional agent constructions. For example, Lezgian marks the non-volitional agent with the adelative case.

42.5.3 Spatial cases used to mark verb satellites with various circumstantial meanings

Cross-linguistically, the *ablative–instrumental* syncretism is particularly common.

Case markers specialized in the expression of temporal meaning (such as Hungarian *-kor*) are not common. Spatial cases are widely used to encode temporal relations, which is commonly viewed as a metaphorical extension of their basic meaning.

Locative of state, allative of transformation, allative of purpose, and ablative of cause, are also relatively common metaphorical extensions of the use of spatial cases.

Basque has a secondary case marker based on allative (*-ra-ko*) for NPs used as verb dependents with a meaning of purpose.

42.5.4 Partitive

Hungarian illustrates the use of ablative to express partitive.

Basque has a partitive marker *-(r)ik* resulting from the specialization of an allomorph of the ablative marker *-tik*.

42.5.5 Comparative

The use of ablative to mark adjective complements expressing comparison is relatively common.

42.5.6 Spatial cases governed by adpositions

It may happen that case markers apt to encode spatial relations without the help of any adposition have also uses in which they are governed by an adposition. (Basque *zehar* 'throughout' combines with NPs in the locative case, Turkish *kadar* 'until' combines with NPs in the allative case, *sonra* 'after' combines with NPs in the ablative case, etc.)

42.5.7 Spatial cases and the syntactic role of noun dependent

Spatial cases are rarely used to mark NPs in the role of noun dependents.

Some languages have 'adjectivizers' systematically used to convert words or phrases typically used as verb satellites (including NPs marked for spatial cases) into noun dependents (Basque *-ko*,[18] Turkish *-ki*, Akhvakh *-se*).

[18] Traditional Basque grammars wrongly identify *-ko* as a 'local genitive' case. In fact, *-ko* attaches to NPs already inflected for case, and has a purely syntactic function of converting verb satellites into

42.6 Non-spatial cases derived from spatial cases

The generalization of the use of adpositions with originally spatial cases is not the only way by which spatial cases can undergo a change in their status. Another possible scenario is that, after developing non-spatial uses, spatial cases tend to be used predominantly in their new function, and to be replaced by other spatial cases (or by adpositions) for the expression of spatial meanings.

This is quite obviously a very common type of change in the evolution of case systems, and, diachronically, syncretisms such as those described in sections 42.4–5 often constitute intermediate stages in evolutions converting spatial cases into markers of core syntactic relations, or into cases expressing other types of circumstantial meanings.

For example, the system of spatial cases of Finnish originally consisted of nine cases organized in three series with the meanings 'unmarked for relative orentation', 'IN' (originally characterized by a formative *-s-*), and 'AT' (originally characterized by a formative *-l-*), but the first series has developed non-spatial uses (modal < locative, partitive < ablative, and transformative < allative), and is now used only residually in spatial functions.[19]

Acknowledgements

I thank the following colleagues for their comments and suggestions: Gilles Authier, Misha Daniel, Bernard Fradin, Sylvain Patri, and an anonymous reviewer.

noun dependents. The source of the error is that, due to morphophonological rules, when *-ko* attaches to NPs in the locative case, the locative suffix may not be apparent.

[19] The three cases of this series are designated in Finnish grammars as essive, partitive, and translative.

CHAPTER 43

THE VOCATIVE – AN OUTLIER CASE

MICHAEL DANIEL
ANDREW SPENCER

43.1 Introduction

The vocative is a form used for calling out and attracting or maintaining the addressee's attention. Unlike some other forms used in this way, such as imperatives (*Listen!* or *Look here!*), a vocative names the addressee explicitly, by using a term referring to and, so to speak, directly acting on them. We will call this function 'address', and any form used in this way, 'form of address'. Forms of address widely attested in the languages of the world are vocative constructions using vocative particles, vocative prosody, and vocative forms of nominals. Sometimes, the form of address is integrated into the case paradigm of the language, and then we can say there is a vocative case. Below we will also refer to the latter as simply 'vocative'. The vocative (case) is thus defined as 'a case used as a form of address, i.e. calling for the addressee's attention by naming them in an explicit way'. It should be stressed that 'address' has to be interpreted rather broadly. This is particularly clear when the vocative-marked name is allowed to come at the end of an utterance rather than at the beginning, as in utterances of the form 'What shall we do now, Mary?' We have found very little literature dealing specifically with vocatives, their forms and functions (but see Jakobson 1936; Zwicky 1974; Xrakovskij and Volodin 1986), though there is a fair deal of work within the field of Conversation Analysis dealing

with various 'vocative' (i.e. form of address) functions (mainly in English, which lacks a case system).

For most theoreticians and descriptive grammarians the vocative is something of an outsider. Many languages have no dedicated case marker and the nominative is used as a form of address. In other languages we see some kind of prosodic process, often unique to the form of address, which is claimed to be 'expressive' rather than 'grammatical'. Grammarians tend to assume that the vocative is primarily used in isolation, which leaves it outside the scope of traditional grammatical investigation. On the other hand, there are a good many languages in which the vocative is a fully fledged member of the case inflection system and even enters into case agreement. Its discourse functions are far from easy to account for.

43.2 Vocative as a morphological case

The strongest notion of 'case' is that in which case is cumulated with other categories and/or is marked differently in different inflectional classes (declensions). It is rare for languages to have a vocative plural form (however, see Georgian, Vogt 1971: 30–1), but several Slavonic languages have a vocative which is sensitive to inflectional class. This even includes Modern Bulgarian which has lost the rest of the case system. Thus, masculine and feminine nouns in the singular have special suffixed forms (Scatton 1984: 140–2, 1993: 199–201), that reflect not only original differences in declension patterns (cf. *sin* 'son' *sine*, *učitel* 'teacher' *učitelju*, *məž* 'man' *məžo*, *žena* 'woman' *ženo*, *carica* 'tsarina' *carice*, *sədija* 'judge' *sədijo*) but also inherited alternations of the final consonant of the stem (cf. *knjaz* 'prince' *knjaže*, *səprug* 'husband' *səpruže*). The fate of the vocative in other languages also indicates the unique status of this case. Old Church Slavonic (often taken as a surrogate for Common Slavonic) had a vocative case. In Czech, Polish, Serbian/Croatian, Upper Sorbian, Kashubian (Cassubian), and Ukrainian the Slavonic vocative is retained. In Slovak, Lower Sorbian, Polabian, Russian, and Byelorussian the original vocative is now lost, except for a few fixed expressions, though Russian has recently developed a 'new vocative' (see below).

This cross-linguistic patterning illustrates the unusual status of the vocative. In general, the more closely related Slavonic languages have similar if not identical case systems, except for the vocative. In Czech the vocative maintains a complex morphology with a variety of allomorphs and various irregular forms, while in the mutually intelligible Slovak it has disappeared. The same can be said of Upper vs. Lower Sorbian and Polish/Cassubian vs. Polabian. A similar situation has arisen in the Indo-Aryan languages. Like Bulgarian, Hindi-Urdu, for instance, has lost most

of the case system inherited from Sanskrit, retaining just a distinction between the 'direct' and 'oblique' forms of a noun stem (distinct in singular and plural forms), and then not for all nouns. However, a separate vocative form has also been retained.

The Russian 'new vocative' (Panov 1997, Comrie et al. 1996, Yadroff 1996) presents an interesting illustration of the way that the vocative can be only partially integrated into an otherwise robust case system. The Proto-Slavonic vocative was lost in the middle ages. The new vocative arose in the twentieth century for nouns whose form is *(. . .)CV'C(C)-a*$_{\text{NOM}}$. This is truncation of the ending, not just the use of a stem or a stem plus zero case marker, because final voiced obstruents remain (or used to remain until recently) voiced, violating an otherwise exceptionless phonotactic constraint (Panov 1997: 108–10): ser'o[ʒ] 'Seryozha', ser'og 'Seryoga' (derivatives of *Sergei*). In addition, the new vocative fails to trigger 'yer vocalization': compare *Serjožk-a*, nominative singular, *Serjožek-ø* genitive plural, with the vocative form *Serjožk*. Such behaviour makes the new vocative unique in the Russian inflectional system.

43.3 Types of marking

43.3.1 No marking

It is fairly common for a language not to have formal marking for a form of address, and to use the nominative instead. Even where a separate vocative form exists, it can readily be substituted by a nominative form in its typical contexts (see below).

43.3.2 Prosody

Prosodic marking, sometimes violating the language's suprasegmental system, is very common with vocatives, with or without other morphological marking such as affixation. Often the suprasegmental feature is not grammaticalized elsewhere in the language (e.g. vowel length in Archi, below). Of course, vocative functions are often associated with specific intonation patterns (such as the so-called 'vocative chant' found in English and other languages), but here we provide a number of examples which are (apparently) independent of intonation.

Stress shift: In Chukchi (Skorik 1961: 306) stress is claimed to shift to the last syllable, as in nominative *kávav* (proper name) – vocative *kaváv* (stress movement is the only signal). In Mangarayi (Merlan 1982: 77) 'The final syllable increases and

then falls sharply in loudness and pitch' (normally the penultimate is stressed). A similar phenomenon is found in Shipibo-Conibo (Faust 1973: 31, 96). Sanskrit has affixal vocative markers but in addition word accent is displaced to the first syllable of the vocative form if the noun is sentence-initial. If the vocative has a premodifier (e.g. adjective or genitive case noun) the premodifier gets the initial accent. Elsewhere in a syntagma the vocative is accentless (Whitney 1879: 108f).

Tone alternation: This is illustrated by Ngiti: *iyamà* 'my mother', vocative *iyamã* (Low-Mid) (Kutsch Lojenga 1994: 166–7).

Vowel lengthening: This is a fairly common phenomenon (e.g. Southern Sierra Miwok, Broadbent 1964: 50; Nivkh, Panfilov 1962: 163). In Archi the last syllable is lengthened, nominative *moHammad* 'personal name', vocative *moHammōd*, even though a long post-tonic syllable is not otherwise allowed word-finally (Kibrik et al. 1977: 295).

Reduction: A common process is phonological (i.e. not morphological) truncation (cf. the Russian 'new vocative'), similar to common types of hypocoristic formation. In Nivkh kinship terms and personal names the final consonant is dropped and the vowel lengthened and/or made more open: *at'ik* 'younger brother' vocative *at'ā*; *T'inyk* (feminine personal name) vocative *T'inē* (Panfilov 1962: 163). In Chukchi masculine personal names in *-tegyn/tagyn* drop the final syllable. Similar phenomena are recorded in Yapese (Jensen 1977: 134). Vogt suggests that Georgian vocatives in *-i* may be derived by truncating vocative/hypocoristic forms in *-ilo* (*dedi* from *dedilo* 'Mom'; *mami* from *mamilo* 'Dad') (Vogt 1971: 20). Further examples are vocative forms of the Georgian second person pronouns (Vogt 1971: 39) or several truncated kinship vocatives in Russian (e.g. *papa* 'Dad', vocative *pa*).

Consonant mutation: In Welsh an earlier vocative prenominal particle triggered soft mutation (voiceless to voiced stop alternation), but the particle has been lost leaving just the mutation (Rhys Jones 1977: 337) *plant* 'children' *bore da, blant!* 'Good day, children!'

Sometimes, an 'expressive vocative' will be accompanied by changes in the segmental composition of the noun. This often looks like straightforward affixation. Examples from languages cited above are Nivkh *-a*, *-e*, Chukchi *-j*, Mangarayi *-j* (optionally).

Most of these types can be thought of as instances of special 'expressive' morphology of the kind often found with diminutives (Zwicky and Pullum 1987).

43.3.3 Case forms proper

We've discussed Slavic vocatives which are fully integrated into the declensional pattern. Other examples are Latin: nominative *lupus* 'wolf' vocative *lupe*, nominative

filius 'son' vocative *filī* 'son'; Ancient Greek: nominative *Ártemis* 'Artemis' vocative *Ártemi*; nominative *patēr* 'father' vocative *páter*; Georgian: nominative *kac'i* 'friend' vocative *kac'o*. Forms of address in Kati (Grjunberg 1980: 178) or Ob-Ugric (Lytkin et al. 1976: 282, 307) could well qualify as case forms (though they are not always considered as such): Kati: *mřor* 'host' *mřóro*.

The Kati vocative is most often homophonous with the oblique plural, while Ob-Ugric (Khanti, Mansi) vocatives are often homophonous with the allative, but from a typological perspective these homonymies seem to be accidental. A more important syncretism that is sometimes observed is with possessed forms. In Udihe the vocative is almost always identical to the first person singular possessed form, but there are slight differences with some words, suggesting that it is, in fact, a separate case (if a marginal one) (Nikolaeva and Tolskaya 2001: 470f). A striking instance of syncretism is found in Scots Gaelic (Mackinnon 1971: 171f). In masculine nouns the vocative is identical to the genitive case but with aspirate mutation (in which a stop alternates with a fricative): *Calum* /kaləm/ 'proper name', genitive *Caluim* /kalimʲ/, vocative *a Chaluim* /əxalimʲ/. With feminine nouns the vocative is the same as the nominative aspirated: *Màiri* /ma:ri:/ 'proper name', vocative *a Mhàiri* /əva:ri:/. In nouns whose nominative plural is syncretic with the genitive singular, the vocative plural is formed by adding *-a* to the aspirated form of the nominative singular: *balach* /baləx/ 'boy', vocative plural *a bhalacha* /əvaləxə/. Otherwise, the vocative plural is the aspirated form of the nominative plural: *caileag* /kalʲak/ 'girl', nominative plural *caileagan* /kalʲakən/, vocative plural *a chaileagan* /əxalʲakən/. (The Gaelic vocative is preceded by a vocative particle *a* which, like many such particles in Celtic, triggers consonant mutation.)

A remarkably robust observation is that the vocative doesn't seem to be marked by a consonant in any language that we know of (other than glides, as *-j* in e.g. Mangarayi; one exception we are aware of is *-v* in Georgian, which is however a contextually motivated, post-vocalic, realization of the main vocative marker *-o*).

43.3.4 Vocative particles

Many languages use a vocative particle, either pre- or more rarely, post-nominally. Often the particle takes an unmarked or a nominative case form of the noun. In many Oceanic languages a particle *e* precedes nouns unmarked by postpositive case particles. The Albanian vocative particle *o* can stand before or after the noun: *O Agim*/*Agimo*! 'Agim!' (Buchholz and Fiedler 1987: 215). However, in other languages the particle governs a marked form of the noun. The Modern Standard Arabic *yaa* precedes the unmarked (nominative) form, but if the addressee noun is in a possessive form (the construct state or pronominal agreement) it takes the accusative: *yaa 'amiir-a-naa* 'O our prince-ACC-1PL!', *yaa 'amiir-a l-muu'miniina*

[O prince-ACC the-believers] 'O Commander of the Faithful!', and if the addressee is non-specific it takes the accusative indefinite form (with 'nunation') *yaa rajul-a-n* [O man-ACC-INDEF] 'someone!' (Abboud and McCarus 1983: 287–8). In Ancient Greek the affixal vocative optionally co-occurs with a prenominal *ō* as in *ō p^{h}il-e* 'O friend!'. In Modern Greek the particles *more*, *re*, or *vre* are used for *ō* (though they are considered impolite addressed to strangers, Holton, Mackridge, and Philippaki-Warburton 1997: 275). See also the Scots Gaelic with mutation-triggering particle *a* discussed in the previous subsection.

43.4 The nominative connection

A dedicated vocative form is often derived from nominative, even if other cases are derived from a different stem. Examples are seen in Figures 43.1, 43.2.

This seems to suggest there is a special connection between nominative (as a syntactic function) and form of address. However, this is probably not the case. In fact, the vocative form tends to be related to the unmarked stem rather than specifically to the nominative, and sometimes is even less marked than nominative. For instance, in Mangarayi the vocative does not take the case/class prefixes which are present in the nominative and other cases (Merlan 1982: 77). A similar process of prefix dropping apparently takes place in Southeastern Pomo (Moshinsky 1974:



Figure 43.1. Chukchi (plural) (Skorik 1961: 308): Nominative and all other cases derived from stem, vocative derived from nominative

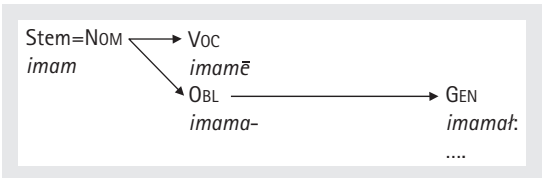


Figure 43.2. Bagvalal (Kibrik 2001: 146): Vocative derived from direct stem (nominative), all other cases derived from oblique stem

103). In Georgian, personal names with a consonantal stem drop the nominative *-i* in the vocative, resulting in a bare stem (Vogt 1971: 19). Some of the Ancient Greek vocatives are bare stems, too, while the nominative forms consist of a stem which has undergone some morphophonological alteration such as vowel lengthening and stress shift, cf. NOM *patēr* ~ VOC *páter* 'father'.

43.5 Functional variation

Probably even less typological research has been devoted to the functional diversity of vocatives than to their formal variation. Some languages distinguish the function of controlling the addressee's attention (discourse vocatives) from simple call-outs. 'Morphological' vocatives are used for attention catching while 'expressive morphology' is used for call outs.

Languages may distinguish several different types of vocative, e.g. Hualapai (Watahomigie et al. 2001: 55–6):

(1) *-é* 'calling the attention of a person who is near the speaker or within the speaker's sight'
John-e mi-yuw-k gwe ma-ma:-'
John-VOC 2-come.here-SS thing 2/3.eat-IMP
'John! Come here and eat!'

(2) *-[w]ó* 'When the person that the speaker is calling for is not in sight'
nya jið-o ge mi-ya:m-ay-ng-yo
1 mother where 2-go-IRREAL-SS.2-AUX/BE.DETACH
'My mother! (I can't see you, but I know you are somewhere here or there.) Where are you going?'

In Chukchi (Skorik 1961: 216–18, 306–9; see also Spencer and Otoguro 2005: 131) second person predicative forms (*tumgy-turi*! 'friends-2PL') are used as 'close range' vocatives, as opposed to the forms in Figure 43.1 (cf. also other Chukotko-Kamchatkan data in Kibrik et al. 2000; Zhukova 1972).

43.6 Lexical domain

The core lexical domain exhibiting vocatives is proper names/kinship terms (including personifications in folk tales), and sometimes including animate

non-humans, usually domestic animals. Presumably, a language which marks vocatives at all will at least mark either kin terms or proper names. Some languages also allow vocatives from inanimate nouns, mostly for rhetoric or stylistic usage; cf. Bulgarian vocative *grade* 'O, city!' or the following Georgian example:

(3) *čem-o sataq'van-o samšobl-o*
my-VOC beloved-VOC fatherland-VOC!
'My beloved fatherland' (Vogt 1971: 20)

Vocative personal pronouns are found in Georgian (Vogt 1971: 39), while vocative demonstratives are attested in Eskimo (Menovshikov and Vakhtin 1990), Nivkh (Panfilov 1962: 243) and Mangarayi (Merlan 1982: 77). For instance, in Nivkh, *tyd'a*, formally a vocative of *tyd'* 'this one', is used as address between husband and wife.

Some languages have appellatives, irregular forms of address, apparently with nominal reference, but related to the respective nominal stem only irregularly or totally unrelated to it. Examples are Avar *le* 'husband (addressed by wife)', *jo* 'wife (addressed by husband)'. These cannot be derived from any other lexical item. Compare the Russian irregular vocatives *pa* 'Dad!', *ma* 'Mom!' which exist alongside regular 'new vocatives' *pap* and *mam*, and *ba* 'grandmom!' (the latter having no regular new vocative counterpart). Appellatives of this sort are akin to interjections.

43.7 Morphosyntax

The vocative is often considered 'extragrammatical', because it doesn't serve to express a grammatical relation (verb argument or verb/noun modifier). However, occasionally it is integrated into the agreement system, especially with possessive adjectives, as in Georgian (4) (Vogt 1971: 19) and Latin (5):

(4) *čem-o k'arg-o*
my-VOC dear- VOC
'my dear'

(5) *fīlī mī*
son.VOC my.VOC
'my son!'

In Old Georgian the vocative even participates in double case marking ('Suffixaufnahme') (Boeder 1995: 159):

(6) *asul-n-o Ierusalem-isa-n-o*
daughter-PL-VOC Jerusalem-GEN-PL-VOC
'You, daughters of Jerusalem!'

Further examples of vocative agreement from Ancient Greek and even Bulgarian are presented in Spencer and Otoguro (2005: 130).

43.8 Peripheral vocative functions

In keeping with their expressive functions, we often find vocatives with non-core discourse functions such as evaluation. Examples (7, 8) are from Georgian:

(7) *še* *oxer-o!*
you.VOC rascal-VOC
'You rascal!'

(8) *tkve* *briq'v-eb-o!*
you.PL.VOC boor-PL-VOC!
'You boors!'

43.9 Conclusions

Languages may have a dedicated vocative form integrated into the case morphology and in a few languages the vocative even makes its way into the morphosyntax and controls agreement. But more often than not the vocative exhibits unusual morphology or phonology, or is not a case form at all. The functions of the vocative distinguish it sharply from 'ordinary' cases, to the extent that some have denied that the vocative can be considered a case at all. The fact that the vocative can form part of the case paradigm without realizing any recognized grammatical or other case-like relation means that it poses an interesting challenge to our conceptions of what a case is.

CHAPTER 44

RARE AND 'EXOTIC' CASES

ANDREJ MALCHUKOV

44.1 Introduction[1]

The present chapter addresses rare phenomena in case-marking, in particular focusing on those which are exceptional in terms of distribution or function. Note that I won't address here unusual formal patterns of case marking, which are briefly discussed in Spencer (Chapter 12; cf. Dryer 2005a). These include such phenomena as prefixal case (found in some Bantu and Berber languages), tonal case (in Afro-Asiatic and Nilo-Saharan languages), and case expressed by stem modification (including cases of mutation, as in Nias). Interestingly, such systems frequently display functional peculiarities, as well. Thus, languages with a tonal case are the same ones that have marked nominatives, a cross-linguistically unusual pattern (König, Chapter 35). Nias is exceptional in that it has marked absolutive and unmarked ergative, in violation of the general markedness pattern (Brown 2001). Finally, the prefixal case in some Bantu languages (e.g. Zulu) may be alternatively analysed as a locative class prefix (Dryer 2005b: 211).

The chapter is organized as follows. Section 44.2 deals with cases deviant in distribution, section 44.3 addresses functionally unusual cases, while section 44.4

[1] Acknowledgements. I am grateful to Jeff Good, Martin Haspelmath, Edith Moravcsik, and Misha Daniel for helpful comments on an earlier version of this paper. I gratefully acknowledge the financial support of the Netherlands Organisation for Scientific Research (NWO) for the research reported here (grant no. 220-70-003 'Case cross-linguistically').

discusses cases deviant in both distribution and function. It will be shown that in most cases the rise of cross-linguistically unusual patterns can be straightforwardly explained in terms of functional and/or diachronic factors.

44.2 Cases with unusual distribution

44.2.1 Multiple case marking

Probably the best studied case of deviant distribution is the phenomenon of **double case** marking or *Suffixaufnahme* (Plank 1995a); see also Spencer, Chapter 12; Moravcsik, Chapter 15). The most wide-spread pattern of *Suffixaufnahme* involves the genitive signalling the dependency within the NP in combination with an external case signalling agreement with the head (see example (1) from Kayardild below). Given that the functions of the two cases are different,[2] double case marking is functionally well motivated and 'logical' (Blake 2001: 103). Yet, most languages impose a restriction on double case marking as case markers tend to 'paradigmaticize' as they get more grammaticalized (Lehmann 1995), and most take recourse to alternative strategies to avoid case doubling (Moravcsik, Chapter 15).

Some of the most spectacular instances of multiple case marking come from Australian languages such as Kayardild (Evans 1995a) and Nyamal (Dench, Chapter 52). Consider (1) illustrating multiple case marking in Kayardild:

Kayardild (Evans 1995a: 115)

(1) *Maku* *yalawu-yarra* *yakuri-na* *dangka-karra-nguni-na* *mijil-nguni-na*
woman catch-PST fish-MABL man-GEN-INS-MABL net-INS-MABL
'The woman caught some fish with the man's net'

In Kayardild, a noun can appear with up to four overt cases markers simultaneously (cf. Dench, Chapter 52, on Nyamal). Dench and Evans (1988) and Evans (1995a) explain this pattern by a rule of 'complete concord' which involves consistent percolation of case from head to dependent. This principle can account for why the possessor takes, in addition to the GEN, the INS case to agree with the possessum; the occurrence of the external 'modal ablative' in this example needs a separate explanation though (see 44.4.2).

Multiple case marking may also arise in languages with so-called 'templatic' morphology, if case markers are distributed across several slots in the template

[2] Indeed, some authors consider these two cases as representative of two distinct categories (cf. case proper vs. concordial case in Mel'čuk's 1998 framework).

(see Spencer 1991 for discussion of 'layered' vs. 'templatic' morphology). Koasati may serve to illustrate this phenomenon, which can be labelled **distributed case.** In this Muskogean language the order of syntactic and spatial cases is different: semantic (locative) cases precede the suffixal determiner ('article'), while syntactic cases (nominative and accusative) follow it (Kimball 1985: 345). This may yield double case-marking, as in the following example:

Koasati (Kimball 1985, ex. 57)

(2) *l·iyá:li-fa-kitt-on (hí:ca-l)*
stepping-LOC-ART-ACC.FOC (see-1SG)
'I see its footprints (lit. place where he stepped)'

The origin of the distributed case pattern becomes evident once one takes into account other functions of case markers in question. On the one hand, locative cases retain some derivational functions (note that in (2) the locative case acts as a sort of nominalizer). On the other hand, syntactic cases seem to have been recently reanalysed from discourse markers: they have special focus forms (as can also be seen in (2)), and are incompatible with other discourse markers (Kimball 1985). Thus the origin of syntactic cases from discourse markers and semantic cases from derivational markers can account for their linear position as well as for their combinatory possibilities.

A pattern reminiscent of double marking is **case layering** as familiar from Indo-Aryan languages where case markers of postpositional origin attach to the oblique form of the noun; cf. in Hindi: *bacce=ne* [child.OBL=ERG], *bacce=ko* [child.OBL=ACC]. While diachronically case layering is a result of renewal of case systems through grammaticalization of postpositions, from a synchronic perspective such patterns are better analysed as either postpositions in combination with the oblique case of a noun (this is the standard analysis of the Hindi forms above), or case proper attaching to an oblique stem (see Kulikov, Chapter 28, for further discussion of case layering in Indo-Aryan; and Spencer, Chapter 12, for a general discussion).

44.2.2 'Head-marking case'

Since Nichols (1986), it is common to speak of head-marking (or 'indexing' in Haspelmath's terms) and dependent-marking (through case or adposition) as alternative strategies for encoding grammatical relations. Yet, as far as formal encoding is concerned, head-marking and dependent-marking generally show almost no formal similarities in languages where both are present. Relatively uncommon are cases such as Sumerian and Abkhaz, where the same or similar case markers may appear either on an argument (as case) or on the verb (as agreement or applicative markers). In Abkhaz, according to Hewitt (1979), for example, some oblique

case markers (e.g. the benefactive *-zə*, the instrumental *-la*) can appear either as postpositions (in (3)) or as 'intraverbal relational particles' (in (4)):

Abkhaz (Hewitt 1979: 113)

(3) *Axra yə-**zə** yə-qá-s-c'e-yt'*
Axra him-**for** it-PREV-I-do-FIN
'I did it for Axra'

(4) *Axra ø-yə-**zə**-q'a-s-c'e-yt'*
Axra (it)-him-**for**-PERV-I-do-FIN
'I did it for Axra'

Similarly, in Sumerian (Hayes 1997; Edzard 2003), verbal 'dimensional' prefixes as in (5) closely correspond in shape to forms of oblique cases:

Sumerian (Hayes 1997: 22)

(5) *mu-**na-ra-ni**-e-eš*
CP-**DAT-ABL-LOC**-go-PL
'They came out from there for him'

The rise of such 'intraverbal' case-markers may be attributed to a particular diachronic scenario where case-marked bound pronouns have been incorporated into the verb form. This origin is still evident in Abkhaz as the case markers invariably attach to a pronominal host. The process of reanalysis is more advanced in Sumerian, where some dimensional prefixes may appear with bound pronouns (e.g. *-b-da-* [it-COM] 'with it'; Hayes 1997: 22),[3] while the others do not. Thus, in both languages, head-marking case appears to have developed from case-markers on incorporated bound pronouns, with a subsequent partial reanalysis of these formatives as applicatives.[4] From a broader perspective, this pattern is related to vestigial case-marking of the sort found in clitic pronouns (e.g. in Romance languages), and – more rarely – on agreement markers (e.g. in the vestigial accusative case on object agreement markers in the Australian language Mangarayi; Merlan 1982).

44.2.3 Displaced case-markers in Tsimshian and Iraqw

In Tsimshian languages case-marking is partially marked by 'connectives', which combine functions of case-markers and determiners. The most peculiar feature of

[3] The dative dimensional prefix *–na-* in (5) is also historically composite, the initial *-n-* being of pronominal origin (Edzard 2003: 92).

[4] Cf. also Haspelmath and Müller-Bardey (2004) for a discussion of a related grammaticalization path leading from adpositions to applicative markers. As the authors point out, the reanalysis starts in the contexts where the object is topicalized and adposition is stranded.

connectives is that they attach phonetically not to the argument whose grammatical function they serve to indicate, but rather to the preceding word. Consider the following examples from (Coast) Tsimshian:

Coast Tsimshian (Dunn 1979: 61 via Iggesen 2005b)

(6) *Yagwa-t niis-da ts'iuu'ts-a laalt*
PROG-3 see-ERG bird-ABS worm
'The bird sees the worm'

In this example, the ERG maker appears in an agreement-like fashion on the verb preceding the agent argument, and the absolutive marker appears on the agent, even though functionally it relates to the following patient argument. Peterson (2006), following earlier suggestions by Boas and Dunn, provides a diachronic scenario for the rise of this peculiar pattern. According to Peterson (2006, and p.c.) connectives developed from a contraction of different elements: agreement markers, case-markers, and determiners. The fact that the pronominal markers are consistently 'left-aligning' must be due to phonological factors: Tsimshian generally does not allow any clitics (on the noun) apart from enclitics (Tyler Peterson, p.c).

Iraqw (Mous 1993) provides another example of a deviant distribution of case, which – in violation of iconicity – attaches to the 'wrong' host. In this Cushitic language some case-markers are suffixes, while some others are enclitics which attach to the noun in preverbal position. Strikingly, the noun 'is not necessarily the object of the case relation' (Mous 1993: 102). Compare the two following synonymous sentences:

Iraqw (Mous 1993: 246)

(7) *inós i hhar-tá hhawat=i hanmiis*
3SG s.3 stick-F.CON man=DIR give
'He is giving a stick to the man'

(8) *inós i hhawatú hhart=i hanmiis*
3SG s.3 man.CON stick=DIR give
'He is giving a stick to the man'

In (7) the object (theme) is marked by the construct state marker (the usual case for objects when they follow a pronominal inflection marker; Mous 1993: 242 and p.c.), while the recipient takes a directional enclitic. In (8), however, the order of the two objects is reversed and the directional case-marker attaches to the theme instead (lit. 'gave man to the stick'). This kind of case displacement arises from interaction of different factors, some of them being syntactic (word order flexibility) and others being morphological (cases are clitics) or morphophonological (clitics appear in a dedicated preverbal position and cliticize to the left). Note that none of these factors is exceptional per se; yet in combination they 'conspire' to produce this unusual pattern.

44.3 Case with unusual functions

44.3.1 Cases with specific functions

Above we considered cases that are deviant in terms of their distribution. The examples to be considered below may be said to be functionally unusual, yet they are unusual in different ways. Note that here we will only consider those functional peculiarities that are substantial and cannot be reduced to matters of terminology (see Haspelmath, Chapter 33, for further discussion). One common explanation that can be given for cross-linguistically unusual cases is that their function is so specific that it is unlikely to often grammaticalize. Not surprisingly, such functionally restricted cases are more likely to be found in languages which have large case systems. For example, in Hungarian, which according to Kenesei et al. (1998) has twenty-seven cases in total, we encounter many cases for which we do not readily find equivalents in other languages (e.g. causal-finalis, modal-essive, distributive, as well as a plethora of local cases). This is expected from the grammaticalization perspective, as embodied in Blake's (2001) inflectional case hierarchy (see Malchukov and Spencer, Chapter 45). More specific notions are less likely to be grammaticalized, and would be expressed in other languages through the use of adpositions or derivationally (note that also for Hungarian other authors regard some of these markers as derivational rather than inflectional).

Similar examples may be cited from many other languages (Haspelmath, Chapter 33; cf. Iggesen 2005b). For example, Shipibo-Konibo (Valenzuela 1997) has a special 'chezative' case for animate locations; many Australian languages have an 'aversive' case ('for fear of X'; Blake 2001), (Icari) Dargwa (Sumbatova and Mutalov 2003) has a dedicated 'contentive' case ('(think/speak) about X'), and a number of other Daghestanian languages have a special 'affective' case for marking experiencers (Daniel and Ganenkov, Chapter 46). Daghestanian languages are also celebrated for having a bewildering number of spatial cases, even if this number may turn out to be exaggerated (Daniel and Ganenkov, Chapter 46; see also see Creissels, Chapter 42, for a general discussion).

Note that for some of these languages availability of unusual cases cannot be seen as an automatic consequence of a large case system: rather case functions are split in unusual ways, with certain uses being singled out by a dedicated case marker. Thus, it is unusual to have a special dedicated case for the 'contentive' ('about') function, which more often than not is assimilated to one of the local functions: approximative (cf. Russian *o*, English *about*), perlative (cf. German *über*, Even *-li*), elative (cf. the 'delative' case in Hungarian). The aversive function is most often expressed by the ablative/source case (as, for example, in Hungarian, Yakut, Armenian, Guaraní, but also in some other Australian languages like Ngiyambaa).

Experiencers are usually encoded through a dative case which is also used to encode beneficiaries and recipients (see Næss, Chapter 38).

While it may be impossible to explain (let alone predict) why certain cases are grammaticalized in individual languages, in some cases this may be related to the 'language profile', that is to the degree of prominence a language attributes to a certain functional domain. Thus, it is hardly surprising that the languages of highland nations such as the Daghestanian peoples of the Caucasus will also display a large number of spatial cases. Similarly, Kiranti languages, spoken in Himalayas, are exceptional in manifesting a phenomenon of 'altitudinal' cases indicating locations higher, lower, or at the same level as the reference point (cf. e.g. in Camling: *kham-dhi* 'to the house (at a higher altitude)'; *khim-i-ni* 'to the house (at a lower altitude)'; *khim-ya-ki* 'from the house (at the same altitude)'; Ebert 1997a: 47).

In other cases, the 'language profile' is determined by sociolinguistic factors, rather than by the geographic setting. Thus, in Korean, which is a 'honorific prominent language', it is not surprising to find such distinctions in the domain of case-marking as well. Thus, four different dative case-markers are found in Korean: the formal dative in *eykey*, the informal dative in *hanthey*, the deferential dative in *kkey*, and the inanimate dative in *ey* (Sohn 1994: 238). Ultimately, an explanation in terms of 'language profile' is not incompatible with a grammaticalization account of the frequency of certain cases: those functional features which are more salient in a particular language are also more frequently encoded and, hence, more likely to grammaticalize.

44.3.2 Syntactically restricted cases

Above we addressed cases which are cross-linguistically rare because they are functionally restricted and therefore less likely to grammaticalize. Other cases may be unusual in being restricted syntactically rather than functionally. In Nivkh (a Siberian isolate), the so-called dative–accusative case (Panfilov 1962) is restricted to marking causees, as well as subjects of clauses marked as indirect speech and subjects of purposive converbs:

Nivkh (Gruzdeva 1998: 19)

(9) *N'-anx-ø n'a-χ pxi-roχ vi-gu-d*
my-sister I-DAT/ACC forest-DIR go-CAUS-FIN
'My (elder) sister let me go to the forest'

In Yakut (a Siberian Turkic language), the genitive case in *-TIN* appears only on stacked possessors. Note that *-TIN* is lacking in the simple possessive construction in (10), but appears on intermediate stacked possessors, as seen in (11):

Yakut (Stachovsky and Menz 1998: 428)

(10) *učuutal jie-te*
teacher house-3SG.POSS
'teacher's house'

(11) *kini aγa-tïn χara-γïn uu-ta*
s/he father-GEN eye-GEN water-3SG.POSS
'his/her father's tears' (lit. 'his/her father's eye's water')

Diachronically, the form in *-TIN* is a remnant of the Common Turkic genitive case, which has been lost elsewhere in Yakut (see e.g. Stachovsky and Menz 1998: 428). Note that the genitive is retained on intermediate possessors where its use is better functionally motivated: in a simplex possessive construction the use of GEN is rendered redundant by the possessive marking on the head. It is plausible that also in Nivkh we are dealing with a residual case.

44.3.3 'Old' cases with diffuse functions

A case may be unusual if its function, rather than being too specific or restricted, is too general or diffuse to be captured neatly in syntactic or semantic terms. As expected, such examples are more typical of minimal case systems, where a single overarching function of an 'oblique' case is often difficult to establish (see Arkadiev, Chapter 47, on two-term systems of the 'distributing' type), yet they may also be found in languages with large or mid-sized case systems. As an example, consider the functions of the 'oblique' case in Kayardild (Evans 1995a: 149). Its general or basic meaning is unclear: it performs some dative-like functions (e.g. marking purpose arguments or object of middle verbs), but otherwise its functions seem semantically disparate. The explanation for this distribution is historical (Evans 1995a: 148–9). The 'oblique' case is the old dative case (still preserved in the genetically related Yukulta), which has been replaced in its central functions (beneficiary, addressee) by emergent ('verbal') cases (see section 44.4.2). Thus we are dealing here with residual functions of a case which do not conform to any recognizable semantic configuration. Another well-known source for semantically disparate functions is (phonologically conditioned) case syncretism. For example, in Ancient Greek the same 'dative' case is used both in the instrumental function (when combined with inanimate nouns) and with the dative function (when combined with animate nouns). This polyfunctionality, hardly attested elsewhere, has been conditioned by a merger of two originally distinct cases in one form (Luraghi 2003: 51). See Baerman (Chapter 14) and Barðdal and Kulikov (Chapter 30), for more discussion of case syncretism.

One also finds other situations where the use of a particular case appears to be idiosyncratic from a synchronic perspective but is well-motivated from a diachronic

perspective. For example, in Ingush, the verb *ladieG-* 'listen' exceptionally takes the ergative–locative pattern, lacking the absolutive argument which is otherwise obligatory in this language. This is due to the fact that historically *ladieG-* originates from the compound *la+dieG* 'ear+put' (see Nichols 1994b: 119). A similar point can be made with respect to subject-experiencer constructions in European languages. As noted by Haspelmath (2001), European languages are typologically unusual in showing a clear preference for subject-experiencer verbs (such as English *like*), rather than object-experiencer verbs (such as German *gefallen* 'like; please'). Haspelmath attributes this predisposition to the fact that many of the subject-experiencer verbs historically derive from verbs denoting a physical action (e.g. *worry* is derived from 'strangle; seize by the throat'; Haspelmath 2001: 79). Thus in this case, as in the previous one, the explanation for idiosyncratic ('quirky') uses should be sought in diachrony, although in the latter case the semantic evolution of the whole construction should be taken into account as well (see also Malchukov 2005 for discussion of pattern inheritance and pattern unification in case-marking).

44.3.4 'Pragmatic cases'

Another reason why it may be difficult to define the function of a case-marker in syntactic or semantic terms is that its basic function belongs rather to the domain of pragmatics. As an example of a case which defies assimilation to a more familiar label consider the 'presentative' case in Samoan (Mosel and Hovdhaugen 1992), which is listed among fifteen prepositional case-markers:

Samoan (Mosel and Hovdhaugen 1992: 500; 143)

(12) *'O le ulugāli'i ma l=a lā fānau*
PRES ART couple and ART=POSS 3DU children
'There was a couple and their children . . . '

(13) *'O le maile sā fasi e le teine*
PRES ART dog PAST hit ERG ART girl
'The dog was hit by the girl'

It is difficult to find a common denominator for the uses of the presentative case in semantic or syntactic terms. It is used to introduce a clause (see (12)), with nominal predicates, as well as fronted NPs (see (13)) (Mosel and Hovdhaugen 1992: 143; 772). In the latter use it is not restricted to any particular syntactic function: it can be used with the object (as in (13)) or with another argument. As also suggested by the case label, the function of the presentative case is basically pragmatic: it introduces rhematic constituents, including contrastive/new topics. Yet, this marker is paradigmatic with other case-markers and therefore should be considered as a case-marker on language-particular grounds.

The case of Samoan is relatively straightforward, as the presentative marker shows few syntactic restrictions in its distribution. In other languages, however, case-markers show a conflation of pragmatic and syntactic information. Thus, the 'nominative' marker *ga* in Japanese is used with rhematic subjects (Ogawa, Chapter 54); the predicate-focus case in Yukaghir is restricted to intransitive subjects and direct objects (Maslova, Chapter 55); the 'subordinative case' suffix *-a* in Nama marks objects and (nominal) predicates (Hagman 1973: 114); the 'background case' in Iraqw marks oblique constituents when topicalized/fronted (Mous 1993: 108); the name of the 'topical non-subject' case in Tariana speaks for itself (Aikhenvald 2003).

Of course it is not unusual that syntactic cases show a correlation with discourse categories, as the well-known correlation between nominative subjects and topical/given information illustrates (Blake 2001: 133). Yet, with respect to 'pragmatic cases' the situation is rather the reverse: here the pragmatic function seems to be basic, while the syntactic functions are derivative. For example, in Tukang Besi, the case-markers *te* and *na* cannot be uniquely associated with any syntactic function without taking into account word order, agreement, and information structure (Donohue, Chapter 53). As aptly demonstrated by Donohue, the tight integration of syntactic and pragmatic information in case-markers leads to a highly intricate pattern of case interpretation in Tukang Besi (as well as in many other Austronesian languages).

44.4 Cases unusual in distribution and function

44.4.1 Function–form mismatches: the designative case in Tungusic

Some cases to be considered below are deviant both in distribution and function, hence they are related to case patterns discussed in sections 44.2–3 above. One of the deviant case patterns concerns mismatches between case forms and case functions, when two forms perform one function, or one case is used to encode two case relations simultaneously. Unambiguous cases of the former type are more difficult to find than may appear. Note that cases of double case marking discussed in section 44.2.1 do not count as an example of this pattern, as the two cases perform two different functions (as in the case of *Suffixaufnahme*), or might be synchronically analysed as a composite case (as in case of 'case

layering').[5] An opposite case, when one form appears to assign two different functions to different arguments can be illustrated by the 'designative' case in Tungusic languages. Consider the following example from Even (North-Tungusic), where the designative case appears on the direct object in combination with possessive markers (which are obligatory with this case in Even).

Even (Malchukov 1995: 9–10 and fn.)

(14) *Hin turki-ga-s emurem*
your sledge-DESG-2SG.POSS brought
'I brought the sledge for you'

In (14), the designative case marks the head of the possessive phrase in the object position, while the beneficiary is encoded as the possessor (and is cross-referenced by possessive suffixes on the head). The unusual feature of this construction is that the possessor is invariably interpreted as a beneficiary. This can be seen through comparing this construction with a parallel construction with the direct object in the accusative case.

Even

(15) *Hin turki-vu-s emurem*
your sledge-ACC-2SG.POSS brought
'I brought your sledge'

The construction in (15) does not impose any particular interpretation of the possessor: it is more likely to be interpreted as source but also allows for other interpretations. In (14), however, the possessor is interpreted invariably as a beneficiary. Therefore, it can be argued that the designative case in (14) assigns two functions to two different NPs: it marks its host as an object while simultaneously assigning the role of beneficiary to its possessor (Malchukov 1995: 9–10).

The origin of the designative construction in Tungusic is unclear. One tempting explanation is to relate the designative case in *-ga-* to the homophonous verb *ga-* 'take'. Indeed, if the designative construction originated from the fusion of the *ga-* in a non-finite form (future participle? supine?)[6] with its object, that would explain why the possessor (the erstwhile subordinate subject) is interpreted as a beneficiary (for example, the source construction for (14) would mean 'I brought (it) for you to take the sledge'; lit. 'for your sledge-taking'). Further, this scenario explains why the designative case obligatorily combines with possessive suffixes, as non-finite verbs take a possessive style agreement. Another possible explanation is to attribute the designative function to reinterpretation of a basic partitive function in the

[5] An example of this pattern, suggested by Edith Moravcsik (p.c.), comes from Hungarian, where the accusative of the third person singular pronoun may be not only *ő-t* 'he/she-ACC' but, most colloquially, also *ő-t-et* 'he/she-ACC-ACC'.

[6] See Malchukov (2007) for further discussion of this development.

context of possessive markers. This explanation is a better fit for the situation in the genealogically related Evenki, where the indefinite accusative/partitive case in *-jA* is interpreted as designative when co-occurring with possessive suffixes (Nedjalkov 1997; see Kittilä and Malchukov, Chapter 36, for exemplification). Interestingly, a similar situation obtains in Eskimo, where the instrumental case is extended to marking indefinite/partitive objects, but also, when combined with possessive suffixes, can be used in the designative function (Fortescue 1984: 87). These two grammaticalization paths are related insofar as the verb 'take' constitutes a typical source for both object markers and instrumental markers cross-linguistically (cf. Heine and Kuteva 2002: 290).

44.4.2 'Quirky' cases in Kayardild

We shall conclude the discussion of rare and exotic cases with the spectacular case of Kayardild, which is celebrated for its unusual case-marking. Apart from the multiple case-marking mentioned in section 44.2.1, a particularly intriguing feature of Kayardild is the existence of 'modal cases', which are used to express tense/aspect/mood features (see also Spencer, Chapter 12; Moravcsik, Chapter 15, for further discussion). Consider the following examples where the 'modal proprietive' case is used to convey future meaning, while the 'modal ablative' case is used to convey the past meaning.

Kayardild (Evans 1995a: 108)

(16) *ngada warra-ja ngarn-kiring-**ku***
I.NOM go-POT beach-ALL-**MPROP**
'I will go to the beach'

(17) *ngada warra-jarra ngarn-kiring-**kina***
I.NOM go-PST beach-ALL-**MABL**
'I went to the beach'

Note that modal cases (such as proprietive and ablative) can be used elsewhere in their normal 'relational' function marking arguments of the verb. In the modal function, however, they appear in addition to (and externally to) relational cases. The diachronic scenario responsible for the rise of modal cases has been described by Dench and Evans (1988) and Evans (1995a; cf. Blake 2001: 108). According to this scenario, these case-markers originally marked non-finite forms in the complement and adverbial function. In accordance with the general rule of 'total concord' operative in Kayardild (see section 44.2.1 above) the cases percolated from the non-finite verb heading the subordinate clause to its arguments (except for the subject argument). Subsequently, the erstwhile subordinate clauses became increasingly used as main clauses (a process called 'insubordination' by Evans), while case forms

on the verb have been fused with the verbal inflection. The net result is that in the course of this development the cases survived on the dependents but not on the verbal head.

Another exceptional phenomenon, which is restricted to Kayardild and a few other Tangkic languages, is the existence of 'verbal cases'. Verbal cases behave like normal cases syntactically (in particular, they also percolate to dependents), but additionally perform a verbalizing function. Consider (18) where the recipient is marked by the 'verbal dative' case, which further takes the imperative inflection, just as the verb does (see Spencer, Chapter 12, for further discussion and exemplification).

Kayardild (Evans 1995a: 336)

(18) *wuu-ja* *wirrin-da* *ngijin-maru-th*
give-IMP money-NOM 1SG-VDAT-IMP
'Give me the money!'

Again this feature is best understood from a diachronic perspective (Evans 1995a: 182–3). Verbal cases originated in noun–verb compounds, which would explain their verbal inflection (thus, verbal dative in (18) stems from *maru.tha* 'put').[7] In itself, verbal origin cannot explain the peculiarity of this pattern; after all, verbs constitute a common source for case-markers cross-linguistically (Blake 2001). What is critical in the Kayardild case is that the full cycle of reanalysis of verbal forms into cases has not been completed. As is usual for grammaticalization processes in general, functional reanalysis is more advanced here (verbal cases behave like cases syntactically in terms of concord), while morphological reanalysis lags behind (the deverbal markers still retain verbal inflections). Thus, both modal and verbal cases are the borderline phenomena attested in incomplete cycles of reanalysis of case markers: in the former case, we are dealing with incomplete reanalysis of the erstwhile cases into the verbal inflection, while in the latter case we are dealing with incomplete reanalysis of the erstwhile verbs into case markers.

44.5 Conclusions

As we have seen, the rise of rare patterns in case-marking usually has transparent functional and/or diachronic motivations. Cases with especially narrow

[7] Interestingly, on Evans' account (1995a: 178–9), the verbal case has residual government properties inherited from the verb: thus, the verbal dative as the original verb 'put' governs both recipient and theme. This is reminiscent of the double role assignment on the part of the designative case of Even, which as was suggested above may also be due to its verbal origin.

978–0–19–920647–6 45-Spencer-ch44 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 648 of 648 May 16, 2008 19:0

or idiosyncratic functions tend to grammaticalize later than other cases cross-linguistically, unless language-particular facts interfere (see section 44.3.1 for discussion of the 'language profile'). A rare pattern may be due to interaction of general functional motivations with a language-specific constraint (for example, both syntactic and morphophonological factors conspire in producing case shift in Iraqw, discussed in section 44.2.3). Rare patterns can also arise from interaction of universal grammaticalization paths with a language specific rule (e.g. rise of modal cases in Kayardild depends on extensive case-percolation). Finally, incomplete grammaticalization cycles can often be implicated in the presence of functionally or distributionally deviant cases (e.g. 'pragmatic cases', not fully syntacticized yet; 'modal cases' being reanalysed into tense/mood markers; and remnant or emergent cases, in general).

PART VII

SKETCHES OF CASE SYSTEMS

CHAPTER 45

TYPOLOGY OF CASE SYSTEMS

PARAMETERS OF VARIATION

ANDREJ MALCHUKOV
ANDREW SPENCER

45.1 Case systems and the case hierarchy

The sizes of case systems vary dramatically, from the minimal (two case) systems (see Arkadiev, Chapter 47), to the large inventories exemplified by Daghestanian (see Daniel and Ganenkov, Chapter 47).[1] An interesting question is whether there are any constraints on the types of possible case systems in the sense that availability of one case implies availability of another. The most concrete proposal of this kind so far is Blake's case hierarchy. Blake (1994; 2001) following earlier insights by Greenberg (1966) and Silverstein (1993 *et passim*) proposed the Case Hierarchy in (1):

(1) NOM > ERG, ACC > GEN > DAT > LOC > INS, ABL > Others

[1] See also Iggesen's WALS chapter (Iggesen 2005a) for an overview of the geographic distribution of large vs. mid-sized vs. small case inventories in the languages of the world.

This is an implicational hierarchy: presence of a case to the right entails availability of case(s) to the left. Blake (2001: 155–60) provides cross-linguistic evidence to support this claim. Thus two-case languages typically have cases covering the core functions (NOM/DIR vs. ACC/OBL), with other functions being expressed, say, by adpositions. Other languages have a three-case system (NOM–ACC–GEN/OBL as in e.g. Semitic), four cases (NOM–ACC–GEN–DAT/OBL, as in German or Icelandic), five cases (NOM–ACC–GEN–DAT–LOC/OBL, as in Latin[2]), six cases (NOM–ACC–GEN–DAT–LOC–INS as in many Slavonic languages, or NOM–ACC–GEN–DAT–LOC–ABL, as in Turkish). In these examples the expansion of the inflectional case system proceeds in accordance with Blake's hierarchy. According to Blake, further extensions of the case systems cannot be predicted (but see below). Another observation made by Blake is that the case lowest on the hierarchy is a kind of elsewhere case which subsumes a variety of lower 'functions'. This is most obvious for the minimal case systems (see Arkadiev, Chapter 47)

Blake is quick to make qualifications to cover apparent counterexamples. First, a language need not have an overt case marker for some cases. This is obviously true for nominative case which is usually unmarked, but it may hold in other cases, when the corresponding function is expressed by alternative means. As noted by Blake, presence of an overt marker for a core argument is frequently obviated through the use of agreement and/or (strict) word order. Similarly a separate genitive might be missing in languages which have head-marking of possessive relations (as, for example, in most Tungusic languages). Another qualification which may account for another group of counterexamples is that a separate case lower on the hierarchy may be missing due to the fact that a higher case has taken over its duty. This is the reason why Blake did not introduce a separate allative case into the hierarchy, as its function is frequently encoded by the (higher) dative or locative cases.

With these qualifications in mind Blake's hierarchy can be shown to hold for a wide variety of case systems, at least as a tendency. Thus, consider the case of Khanty (Ostyak), which on the surface does not comply with the hierarchy. Khanty (Nikolaeva 1999), has been described as having a four-term system, including nominative, locative, prolative ('along X'), and essive ('as an X'), cases. Putting aside the prolative and essive cases which go beyond the hierarchy, the case system apparently violates the hierarchy as it lacks accusative, genitive, and dative cases, though these should be entailed by the existence of LOC. Yet this pattern can readily be explained within Blake's approach. First, ACC marking is obviated by object agreement; the same holds for dative, since Ostyak makes a distinction between primary and secondary objects rather than direct and indirect objects. Further, it lacks a genitive because the possessive relation is shown by head-marking possessive suffixes. The

[2] In Latin, the lowest case is traditionally referred to as 'ablative', but it encodes the locative (and instrumental) function as well (Blake 2001: 157).

lowest case subsumes a number of different functions which include locative and instrumental.

Nonetheless, there seem to remain genuine counterexamples to the hierarchy, although such patterns are rare. For example, (Modern) Irish has retained only two cases, NOM and GEN (apart from petrified dative and vocative forms in certain declensions), but has no separate ACC case, without (object) agreement or (VS) word order ostensibly compensating for that.[3] Further, there are certain conceptual problems in the upper part of the hierarchy. There is a certain circularity to the claim that nominative is the highest ranking case obligatorily found in any system, when it can only be identified by reference to the lower cases (accusative or ergative) with which it contrasts.[4] Another problematic point concerns Blake's interpretation of the lowest case as an elsewhere case, which subsumes a variety of lower functions. Effectively, this means that the hierarchy restricts case polysemy (syncretism) to adjacent ranks in the lower part, thus coming close to an idea of a semantic map (see Malchukov and Narrog, Chapter 34). Yet this interpretation of the hierarchy is dubious, because the case hierarchy cannot account for all recurrent polysemy patterns. Thus, ACC and DAT are not adjacent on the map even though they are quite often syncretic, and the syncretism need not be mediated by the genitive function as the hierarchy would predict. This deficiency is not unique to the case hierarchy, but holds for one-dimensional models (semantic maps) in general (see Malchukov and Narrog, Chapter 34). Yet, on the whole, Blake's hierarchy correctly captures the general markedness pattern among cases, as well as eligibility of individual cases for grammaticalization. Thus, cases on the lower end are more frequently expressed by case or adposition than those higher on the hierarchy, which may remain unmarked. Marked nominative languages, discussed by König (Chapter 35) represent a counterexample. Another exception is Koasati, where dative (which is also used for possessors and prepositional objects) is zero, while cases higher on the hierarchy (NOM, ACC) are overt. Such systems are cross-linguistically rare, however. On this broad interpretation Blake's hierarchy reflects a clear tendency, and is consistent with the findings of grammaticalization research (Lehmann 1995; Heine, Chapter 29 *et passim*). It is also consistent with hierarchies of case marking as established in the functional-typological literature, although the latter are stated in terms of grammatical relations rather than case. Thus, Lehmann (1988) and Whaley (1997) note that case and agreement interact with the grammatical function

[3] As noted by D. Creissels (p.c.), locative forms of nouns in some other Native American languages (e.g. Iroquian, Uto-Aztekan) can be problematic for the hierarchy, given that there is no case marking of core arguments. Yet, it seems that in most cases these forms are idiosyncratic and do not qualify as cases; furthermore absence of case on (core) arguments is compensated by the system of cross-referencing on the verb.

[4] Yet, as observed by B. Blake (p.c.), the highest position of NOM can be established on the basis of other evidence, such as markedness considerations. In a similar vein, it has also been suggested by Creissels (p.c.) that marked nominatives occupy a lower rank on the hierarchy, as they entail availability of the unmarked form.

hierarchy in opposite ways, case/adposition-marking preferentially marking lower functions, agreement the higher functions:

(2) Relationship between case and agreement (Whaley 1997: 154)
Subject > Direct object > Indirect object > Other
Agreement ⟶
⟵ Case/adposition

Markedness relations captured by Blake's hierarchy may be manifested in other ways as well. Thus, they constrain derivational relations between individual cases: cases lower on the hierarchy may be derived from higher ones but not the other way round. Thus, ABL can be built on LOC as happens in a number of African languages (see an example from Maale below), but the opposite pattern is not found. Similarly in many languages with a large case inventory higher cases serve as the morphological basis for creating further case forms (Spencer, Chapter 12; Blevins, Chapter 13), although when this happens the original case will often get restructured as a stem form whose default function is to express that basic case. Thus, in Daghestanian languages such as Dargi lower cases are built on either ERG or GEN stems (Sumbatova and Mutalov 2003; see Daniel and Ganenkov, Chapter 46).[5] As a final illustration, consider triple case marking in Maale (König 2006 citing Amha 2001), where ACC functions as a stem formative, and LOC serves as a basis for formation of the ablative: *maa-o-idda-ppa* [house-ACC-LOC-ABL] 'from the house'.

Another interpretation of Case Hierarchy pertains to syntagmatic relations between individual cases. On this interpretation availability of a lower (more marked) case in a case frame presupposes availability of a higher (less marked) case. For example, some languages (like Dravidian, mentioned by Blake 2001: 90) allow dative arguments only with transitive verbs, that is, availability of dative in a case frame entails availability of the 'higher' accusative case. Another instructive case comes from Garo (a Tibeto-Burman language; Burling 2004), where dative marking on Recipient switches to accusative, when the theme is omitted. Compare *Ang-a nang·-na gol-po-ka a-gan-a* [I you-DAT story-ACC tell] 'I tell you a story'; *Ang-a nang·-ko a-gan-a* [I you-ACC tell] 'I tell you/speak to you' (Burling 2004: 187). This account (in the form of a case markedness hierarchy) has been implemented in a number of formal proposals, including the case-in-tiers approach (see Maling, Chapter 5 *et passim*), and in OT by Woolford (2001), although they do not explicitly refer to Blake's Case Hierarchy. Note that this is a more specific interpretation of

[5] It should be noted however that the generalized notion of markedness as embodied in Blake's hierarchy is not always helpful, since individual cases may show relations of local markedness with respect to specific noun classes (this affects also derivational markedness, as discussed by Aristar 1997). A good example comes from Koasati which has seven cases, but any given noun may take no more than five cases.

the hierarchy as it makes the claim that a higher case will be found in the same construction, not just in the case inventory.

Switching to a diachronic perspective it has been mentioned that Blake's hierarchy is designed to show that 'morphological case systems grow and decay in a certain order' (Blake 2001: 155), that is, the case hierarchy predicts not only the order in which cases are acquired but also the order in which they are lost (although case loss is not explicitly addressed by Blake).[6] Thus, in South Slavic dialects (Sobolev, Chapter 49) and Ancient Greek (Luraghi 2004b) loss of case proceeds roughly in accordance with the hierarchy. Old French provides a nice illustration too, with just two cases, the so-called 'cas-sujet' conflating Latin NOM and VOC, and the so-called 'cas-régime' conflating ACC, GEN, DAT, and ABL–LOC–INS (Denis Creissels, p.c.). The same seems to be true of mergers of individual cases in different branches of Indo-European (cf. Barðdal and Kulikov, Chapter 30). As is clear from Barðdal and Kulikov's presentation the mergers usually target adjacent ranks on the Case Hierarchy: cf. the NOM-ACC merger in Balkan languages, the DAT-LOC merger in Greek, the LOC-ABL-INS merger in Latin, the DAT-LOC-ABL-INS merger in Celtic and Germanic. In language attrition, too, it seems that the hierarchy is at work. For example, in the speech of Russian-speaking emigrants loss of case generally seems to comply with the hierarchy, even though there is a great amount of variation (Sussex 1993: 1016). The effects of case hierarchy – conceived in general terms – have also been noted in different psycholinguistic domains, as reviewed in Part IV. Thus, cases higher on the hierarchy are acquired earlier as compared to the lower ones (e.g. accusative before dative in German; cf. Eisenbeiss, Narasimhan, and Voeikova, Chapter 24). Similarly, speech errors under language impairment indicate that lexical cases, which are lower on the hierarchy, are more problematic than structural cases (Lamers and Ruigendijk, Chapter 27). And in language comprehension violations resulting from a substitution of higher case for the lower one are felt as less severe as compared to violations of the opposite kind (Bader and Lamers, Chapter 26).

Ultimately Case Hierarchy may be motivated by the frequency of individual cases (see Haspelmath 2006 on the role of frequency in evolution of grammatical systems). Although comprehensive data on frequencies of individual cases is usually lacking, whenever available it seems to be consistent with the hierarchy. Thus in discussion of case loss in Ancient Greek Luraghi (2004b: 366) notes that frequencies of individual cases (in Mycenean Greek) are in accordance with the hierarchy NOM,

[6] This issue has been recently taken up by Hawkins (2004) who derives markedness hierarchies proposed by Greenberg (in particular, his Case Hierarchy, NOM > ACC > DAT > Other), from the general principle 'Minimize Forms' (MiF). Hawkins argues that MiF has a processing explanation ultimately rooted in frequency. Importantly in the present context, one of the consequences of the MiF are the predictions concerning the morphological inventory of forms, which capture the effects of Blake's hierarchy. Hawkins illustrates the effects of the case hierarchy with the gradual reduction of case system in the history of Germanic languages, particularly English.

ACC > GEN > DAT > INS, reflecting the order of case loss. Similarly, it has been observed for Germanic languages that the loss of particular uses of individual cases ('case constructions') depends on their (token) frequency (Barðdal and Kulikov, Chapter 30).

As noted above, Blake was reluctant to pursue the hierarchy beyond the ABL/INS, where no definite ranking can be established. Yet it seems that in a number of specific domains further implicational relations can be established. Thus, Levinson et al. (2003) have suggested a hierarchy for the domain of spatial adpositions; availability of an adposition to the right is said to predict the availability of adpositions to the left:

(3) Scale of topological adpositional relations (Levinson et al. 2003).
AT < IN < ON/UNDER < OVER/NEAR < ON-TOP < ATTACHED < INSIDE ...

Indeed if a language has just one adposition marker it is likely to have a general location ('contact') as one of its functions. If the system is built up further it will differentiate between interior and surface location, etc. An important difference between the case hierarchy and the spatial hierarchy proposed by Levinson et al. is that the latter is intended to represent markedness relations within the system. Thus Levinson et al. note that their hierarchy is not meaning-preserving. This is because the meaning of a 'higher' locative adposition changes with the addition of the lower one, so that the higher and lower adpositions are opposed as unmarked and marked member in a privative opposition. It remains to be seen if Blake's hierarchy can be reconstructed in this way: it seems reasonable for the upper part of the hierarchy, which shows a clear markedness pattern. Thus, as mentioned earlier the nominative shows a similar dependency with respect to ergative and absolutive markers. The only difference here is that core arguments are normally unmarked, unlike locative adpositions.

45.2 Morphological vs. syntactic case

Another general aspect of case systems which is subject to cross-linguistic variation is the relation between morphological vs. syntactic case (cf. Spencer, Chapter 12). By 'morphological case' (m-case) we mean an (inflectional) case form of a nominal, what we might call a formal characterization. On the other hand, 's(yntactic)-case' refers to the case function borne by a NP in a phrase, and this is defined distributionally, in terms of grammatical relations, subcategorization, agreement, and so on. It's important to bear in mind that we are thinking of 'syntactic case' in a descriptive or pre-theoretic sense. If nouns in a language have several (morphological)

cases and adjectives agree with nouns in case then the syntax will of necessity appeal to some feature [Case] in order to capture the generalization that the adjective is marked with the same case as the head noun, that is, somewhere in the grammar we will need the statement that if the noun has case value [Case: α] then the adjective will have [Case: α]. This will be true however the adjective case marking is analysed (as true agreement, as multiple marking, or as something else). Similarly, we may find that all verbs or prepositions of a certain semantic or morphosyntactic class select complements in the same case.

Ideally, the two notions coincide and the morphological marking mirrors syntactic case assignment. In such a system we have a direct mapping between m-case and s-case (and we can safely use the same labels for each type). This can even be true in languages with 'case stacking' (see Dench, Chapter 52; Spencer, Chapter 12), where the stacking is 'iconic', that is, where the order of case affixes on a given word corresponds to the level of embedding of the dependencies, as, for instance, if a genitive-marked noun agrees in, say, allative case with its head, to give a phrase of the form Noun1-GEN-ALL Noun2-ALL. In languages with extensive case stacking, the case marking is always very transparent (agglutinative), and it's not clear that there's actually any need to make reference to morphological case features as such in an explicit grammar.

However, it's actually quite difficult to find languages in which the morphological ('formal') definition and the syntactic ('distributional') definition of case line up perfectly. A good deal of typological attention has been devoted to the issue of when to apply 'formal' as opposed to 'distributional' definitions of case to capture regularities in grammars, particularly in the context of Australian noun/pronoun systems (cf. contributions to Plank 1991a). As is well-known, in many of these languages the lexical nouns exhibit ergative–absolutive patterning, while the pronouns show nominative–accusative patterning. On one approach, adopted by Silverstein and much of the Australianist literature, such systems are described as split ergative: the noun forms have an ergative–absolutive distribution, while the pronominals have nominative–accusative distribution, and so these are the cases those word forms are 'in'. An alternative approach, first proposed in the context of Australian languages by Goddard (1982), but traditional for the description of European languages, treats a discrepancy between nouns and pronouns in terms of case syncretism/allomorphy. On Goddard's approach we assume three cases, ergative, nominative–absolutive, and accusative. In nouns the NOM-ABS case syncretizes with the accusative, while in pronouns the ergative and the NOM-ABS cases syncretize. Hence, there are three s-cases (defined distributionally), although for any given nominal there are only two m-case forms. On Goddard's approach the relation between m-cases and s-cases is complicated. On the split-ergativity approach the relation between case marking and syntax is complicated.

This controversy, frequently referred to under the heading of distributional vs. form-based approaches to identification of cases (Blake 2001), remains unresolved

(see also Iggesen, Chapter 16, and Baerman, Chapter 14, for somewhat different approaches). Blake (2001: 19–25) is non-committal in that respect, noting that the two approaches derive partly from different traditions and partly from the differences in the data addressed: for inflecting languages (Latin, Russian) which have numerous paradigms with case syncretisms and where each case is used in a number of functions, we can only capture the distributional facts by appeal to s-cases ('distributional' account). On the other hand, if a language has case markers which are morphologically very regular, with no inflectional classes, then it will often be possible, in principle, at least, to state regularities over the form classes themselves ('formal' account). For instance, suppose a language has ergative marker *-du* for nouns, accusative marker *-nya* for first/second person pronouns, and leaves absolutive (noun) and nominative (pronoun) forms unmarked. Then we can say that the A (transitive subject) function is expressed by *-du* (nouns) or unmarked (pronouns), the P or O (direct object) function is unmarked (nouns) or expressed by *-nya* (pronouns) and the S (intransitive subject) function is unmarked for all nominals. If the system really is this simple there is then no need for an explicit grammar to refer to any case labels at all (though grammarians will no doubt want to think of *-du* as 'ergative' and *-nya* as 'accusative'). This is a consequence of Beard's Criterion (Spencer, Chapter 12), which essentially says that a [Case] feature is only needed in a grammar in order to generalize over distinct case allomorphs (for instance in different inflectional classes) or to generalize syntactic rules (for instance, adjective–noun case agreement).

In some instances there seems to be no alternative to the distributional approach. A case in point is that of the Russian 'animate accusative' (Spencer, Chapter 12; see also Baerman, Chapter 14). Nouns of the *-a* declension class are mainly feminine and have distinct nominative, accusative, genitive singular forms, e.g. *Maša* 'Mary (dim.)': *Maša*, *Mašu*, *Maši*. Masculine gender nouns of the default masculine consonant-final declension class such as *Aleksandr* 'Alexander' have an accusative which is identical to the genitive, *Aleksandra*. Non-animates such as *skafandr* 'space suit' have an accusative identical to the nominative (the genitive is *skafandra*). Adjectives agreeing with these nouns show the same case syncretisms: *mojego Aleksandra* 'my Alexander (ACC/GEN)', vs. *mojego skafandra* 'my space suit (GEN. only)'. Some masculine animate nouns belong to the *-a* declension class such as the diminutive of *Aleksandr*, namely, *Saša*. These decline like feminines: *Saša, Sašu, Saši*, but they take masculine agreements: *mojego Saši* 'my Sasha (GEN)', but also *mojego Sašu* (ACC.). It's impossible to describe such a patterning without saying that objects are in the s-case ACCUSATIVE which, however, maps in a rather complex fashion onto nominative/accusative/genitive m-case.

Although the m-case/s-case distinction is clear in itself, there are a good many constructions which illustrate the difficulty of deciding whether a given alternation should be attributed to alternations of form (that is, m-case) or alternations in the way syntactic cases are specified or assigned. The largest source of

such analytical indeterminacy comes, perhaps, from differential case marking, for example, differential object marking (DOM). DOM is found when a direct object is sometimes marked by an overt accusative marker and sometimes left unmarked, so that the form systematically syncretizes with the nominative form (and in Finnic languages we may get the added complication of an alternation with a partitive case form; see Ackerman and Moore, 2001, for detailed discussion). The question is then whether the unmarked object is 'really' in the NOM s-case (which we can call the uniform NOM analysis) or whether we have an unmarked allomorph of the ACC s-case which just happens to be homophonous with the nominative m-case form (the unmarked accusative analysis). This is not necessarily quite the same situation as the genitive–accusative in Russian. In Russian there are uncontroversially accusative-marked masculine nouns which nonetheless trigger genitive case agreements, so we know that we are dealing with s-case ACC. In the absence of agreement and other such evidence, however, there remains an indeterminacy in the description.

Here, we consider characteristic examples from Tungusic languages. Like Turkic languages (see e.g. Johanson 2006 *et passim*), Tungusic languages display a phenomenon of differential object marking (DOM). However, unlike Turkic, the object cannot be left unmarked unless it is marked by a reflexive marker (see also Kittilä and Malchukov, Chapter 36). Thus, while the normal way of marking the direct object in the North-Tungusic Even would be with the accusative suffix *-w*, if the object takes a possessive–reflexive suffix (*-j* for the singular possessor and *-vur* for the plural possessor), the accusative marker is lacking:

Even (Malchukov 1995 and fn.)

(4) *d'u-v/d'u-j* *itten*
house-ACC/house-REFL saw
'(He) saw the house/his own house'

On the face of it, absence of overt accusative case in the context of the reflexive marker could be interpreted either as unmarked nominative or as a zero accusative case. The first view is supported by the fact that the reflexive form may also appear in those positions where a nominative rather than an accusative form would be appropriate, for example in the possessor position:

(5) *d'u-j/d'u* *urke-ve-n* *itten*
house-REFL/house door saw
'(He) saw the door of his own house'

However, the attribute modifying the reflexive object must be in the accusative case, which would suggest that REFL is actually an unmarked accusative (or that accusative surfaces as zero in the context of reflexive).

(6) *anngamat-v* *d'u-j* *itten*
new-ACC house-REFL saw
'He saw his new house'

This makes the Tungusic case rather similar to the Russian genitive–accusative, and so it really does look as though we have here an instance of DOM which requires us to mark the object with ACC s-case while giving it the m-case form of the nominative.

45.3 Typology of case paradigms

It is a common observation in the literature that cases mark different sorts of information: syntactic, semantic, and discourse-pragmatic (Givón 1985, 2001;[7] Blake 2001; Kibrik 1997). The distinction between syntactic and semantic cases is, of course, familiar from descriptions of European languages, where it has been alternatively described in terms of grammatical vs. concrete cases (traditional grammar; see Blake, Chapter 1; Haspelmath, Chapter 33), or in terms of structural vs. inherent case (generative grammar; see Butt, Chapter 4). Conflation between syntactic/semantic and discourse-pragmatic information is less conspicuous in European languages, although it is common to regard subject case as a grammaticalized topic and correspondingly relate nominative case to topicality (Givón 1984). Other languages conflate pragmatic functions with grammatical functions in a clearer fashion. Thus in Korean, both nominative and accusative case perform pragmatic (information structural) functions as can be seen in contexts of case stacking and case spreading (Van Valin, Chapter 7). Another well-known example is Japanese, where the topic marker *wa* is to some extent in a paradigmatic distribution with respect to other cases: it excludes nominative (*ga*), accusative (*o*), and dative (*ni*) although it may co-occur with oblique cases such as *de* 'in' and locative uses of *ni* (Ogawa, Chapter 54). But it's not obvious that we should regard *wa* as a case, just by virtue of the fact that it's in a paradigmatic relation with other things commonly called case markers, since there are several other unmistakably discourse markers like *mo* which behave in a similar fashion (see Spencer and Otoguro 2005). Here it's worth noticing that Spencer and Otoguro (2005) point out that none of the Japanese 'case' markers meet Beard's Criterion (see also Spencer, Chapter 12).

Where semantic and syntactic ('structural') cases are concerned it is more common cross-linguistically to have both types within the same paradigm. A possible exception in this respect are languages with a small case system where the lowest case cannot be associated with a particular semantic function (see Arkadiev,

[7] Givón (2001: 201 ff.) distinguishes case marking which is semantically oriented (active–stative languages), syntactically oriented (nominative–accusative languages), and transitivity oriented (ergative–absolutive languages). Earlier Givón (1985) spoke about the 'case-recoverability' problem posed by inflectional case which either encodes semantic or pragmatic information.

978–0–19–920647–6 46-Spencer-ch45 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 661 of 667 May 16, 2008 19:0

Chapter 47). Equally rare are systems where cases higher on the hierarchy qualify as semantic. One example discussed in the literature is Manipuri (Bhatt and Ningomba 1997), where both core cases have been described as basically semantic, even though it is not clear from the terminology. For example, the presence of the agentive case *-nə* (called 'nominative' by Bhatt and Ningomba) in Manipuri signals volitionality, as illustrated by the following examples:

Manipuri (Bhatt and Ningomba 1997: 104)

(7) *.əy-nə tebəl-də theŋŋi.*
I-ERG table-LOC touched
'I touched the table (volitionally).'

(8) *əy tebəl-də theŋŋi.*
I table-LOC touched
'I touched the table (involuntarily).'

Similarly, the use of patientive ('accusative') case *-bu* correlates with affectedness; it is therefore mostly restricted to animate objects, but may also carry over to intransitive patientive subjects (Bhatt and Ningomba 1997: 112). Other possible candidates are languages with 'active' alignment (also known as 'semantic alignment', Donohue and Wichmann 2007), although it manifests itself more often in agreement than case. Some authors also admit the possibility of accusative and ergative systems being organized on a semantic basis. For example, according to Kibrik (1997) Daghestanian languages are semantically based and distinguish semantic hyper-roles of Agentive and Absolutive. Yet this claim should be qualified given that there is a markedness asymmetry between the two cases.[8] That is, while ergative undeniably performs a semantic (agentive) function, interpretation of absolutive in semantic terms is less straightforward (which is not surprising given that the absolutive is morphologically unmarked). This is also consistent with Blake's view, which characterizes ergative as a basically semantic case, nominative as a syntactic case with a pragmatic basis (topic), while accusative and absolutive are syntactic cases with Patient as a prototype. More generally, it is probably true that 'mixed' languages conflating different dimensions are more widespread than 'pure' languages, which are either role-oriented or discourse-oriented (cf. Kibrik 1997).

Commonly, the distinction between semantic and syntactic cases is not clear cut since the same case may encode both sorts of information (cf. Butt 2006: 142–7). For the same reason, Mel'čuk (1998: 328) suggests we speak of semantic vs. syntactic use of cases rather than a semantic vs. syntactic case dichotomy. Thus, although structural cases may seem to be devoid of semantic function this function reappears with

[8] It should be noted though that Kibrik's (1997) distinction between role-oriented vs. flow-oriented strategies, like the related distinction between role-dominated vs. reference-dominated languages of Foley and Van Valin 1984, is based on syntactic argumentation, rather than exclusively on the semantics of case-markers.

978–0–19–920647–6 46-Spencer-ch45 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 662 of 667 May 16, 2008 19:0

case alternations (NOM alternations are indicative of loss of Proto-Agent, while ACC-alternations are indicative of loss of Proto-Patient properties; Ackerman and Moore 2001; de Hoop and Malchukov 2007). De Hoop and Malchukov (2007) attribute semantic effects in alternations of the fluid type to a general constraint against having synonymous forms (see Malchukov and de Swart, Chapter 22, on split vs. fluid alternations). Viewed diachronically this conflation is natural since syntactic cases develop on the one hand from semantic cases through well-known routes of grammaticalization (see references in Heine, Chapter 29; Malchukov and Narrog, Chapter 34). Interestingly, the opposite development (semanticization of the erstwhile syntactic case) is also attested. The best known example is the reanalysis of the ergative to agentive marker as it spreads to (agentive) intransitives, as it happened in Kartvelian and some Iranian Pamir languages (Plank 1985; Harris and Campbell 1995). Given these diachronic connections the distinction between languages with semantic (active) alignment and languages with a syntactic (accusative and ergative) alignment is hard to draw; indeed as shown by Bickel and Nichols (Chapter 20; cf. Merlan 1985; Nichols 1992) there is a continuum between ergative and accusative languages with accusative-based and ergative-based active languages in between.

Another continuum leads from syntactic to pragmatic cases. In familiar Indo-European accusative languages it is common to treat the subject case (nominative) as a conflation of syntactic and discourse features (topicality). The same may be true for accusative case especially in languages allowing object shift (cf. Dryer 1986 who attributes the rise of primary/secondary object distinctions to the fact that P and R arguments are both eligible as secondary topics). This conflation is still more evident in a number of other languages discussed in this volume, Japanese, Yukaghir, Tukang Besi. As is well known, the distinction between the two different 'subject cases' *ga* and *wa* in Japanese has a discourse basis (see Ogawa, Chapter 54, for discussion). Yukaghir has a special 'predicative' case which appears on 'rhematic' (new, focal) P and S arguments (Maslova, Chapter 55). The case of Tukang Besi, as described by Donohue (Chapter 53) is particularly instructive, inasmuch as it shows a tight integration of pragmatic and syntactic information within one system, which leads to a very intricate system of case assignment. Thus 'core' case in *te* is interpreted as P if postverbal and does not control agreement (see (9)); if the verb is bipersonal it is interpreted as A (see (10)), unless it is fronted, in which case it is again interpreted as P (see (11)).

Tukang Besi (Donohue, Chapter 53, exx. (4), (7a), (8))

(9) *Te wowine no-'ita te kadadi.*
TE woman she:saw TE bird
'The woman watched birds.'

(10) *No-'ita='e te wowine na kadadi.*
she:saw:them TE woman NA bird
'The woman watched the birds.'

(11) *Te kadadi no-'ita='e te wowine.*
TE bird she:saw:them TE woman
'The woman watched the birds.'

Clearly, it is futile to try to interpret these constructions in terms of semantic roles; but it is also not straightforward to assign them a syntactic function, since such a characterization would have to make reference to word order (and agreement). The syntactic characterization offered by Donohue is as follows (Donohue, Chapter 53, ex. (19)):

(12) *te*: if preverbal, *te* marks the subject;
if postverbal, *te* marks a core non-subject ('object').
na: necessarily postverbal, *na* marks the subject.

An alternative interpretation, which is in terms of discourse functions, is not totally straightforward either (Donohue, Chapter 53, ex. (20)):

(13) *te*: if preverbal, *te* marks identificational focus;
if postverbal, *te* marks a non-given term.
na: necessarily postverbal, *na* marks a (non-focused) given term.

It seems that part of the problem is that reanalysis of pragmatic to syntactic markers is still under way. However, the pragmatic function is not totally clear either, which suggests intermediate stages of reanalysis. This seems to be true of other languages as well, insofar as the authors discuss the diachronic dimension. Thus, for Yukaghir it has been suggested that P-markers developed from Focus-markers but retain some residual discourse functions (Maslova, Chapter 55). The same is true for languages with the marked nominative case which frequently originates from definiteness/topicality markers (König, Chapter 35; König 2006). Similarly, ERG marking frequently arises from contrastive topicalization (Malchukov 2008a; cf. McGregor 1998). This is exemplified below for Newari, where As take the ergative marker when rhematic-focal (Givón 1984: 154). Thus, (14) would be appropriate as an answer to 'Who is breaking the window?', (15) as an answer to 'What is the man doing?' (ibid.):

Newari (Givón 1984: 154)

(14) *Wō manu-nā ihya tajua-na co-na*
the man-ERG window break-AUX be-AUX
'**The man** is breaking the window'

(15) *Wō manu ihya tajua-na co-na*
the man window break-AUX be-AUX
'The man **is breaking the window**'

More generally languages which have distinguishing case or global case of the type found in Papuan languages are likely to have developed it from pragmatic markers. As we have said, part of the difficulty in assigning them a particular function is that

they present different stages of reanalysis. The mechanism of reanalysis seems to involve conventionalization of implicatures (see Hopper and Traugott 1993 on the role of this factor in semantic evolution). Thus since A-arguments are usually given and P-arguments new, definiteness/topicality markers might be interpreted as an A-case and rhematic/focus markers as a P-case.

45.4 Case-marking: beyond morphological case

It is conventional to regard case on a par with alternative strategies such as agreement and word order (cf. Kiparsky's view of case, agreement, and word order as alternative linkers; see Butt, Chapter 4). This approach makes sense in particular with regard to distinguishing case as found in Papuan languages which shows an interaction between these mechanisms for purposes of argument differentiation.[9] Thus case-marking in Fore may be obviated through the use of agreement (see Malchukov and De Swart, Chapter 22). If case marking is not possible, or is not unambiguous, for particular words or constructions, we may find the phenomenon of word order freezing (or conversely we may find that case-marking allows scrambling; see Neeleman and Weerman, Chapter 18). This is also related to Blake's view noted above that the absence of particular case can be compensated through the use of agreement and word order.

Yet the parallels between these mechanisms are only partial (see Siewierska and Bakker, Chapter 19). Thus, as a means of disambiguation word order can be reasonably compared to core cases, but not to oblique cases. Even for core cases the parallels are incomplete. Postverbal position can allow for a much broader range of interpretations compared to accusative case. The postverbal NP cannot be generally identified with the direct object as illustrate the following examples from Lao (Enfield, Chapter 57; see the characterization of the case frames provided for these examples by Enfield):

Lao (Enfield, Chapter 57, exx.24, 25)

(16) *khòòj*^{5} *nak*^{2} *kapaw*^{3} *nii*^{4}
I heavy bag DEM
'I find this bag heavy.' (Experiencer subject)

[9] It should be kept in mind though that other sorts of information may be used for disambiguation as well; see de Hoop and Lamers (2006) on distinguishability constraints, which apart from case, agreement, and word order pertain to selectional restrictions, animacy information, and the role of context.

(17) *kapaw*[3] *nii*[4] *nak*[2] *law*[5]
bag DEM heavy liquor
'The bag is heavy from the liquor (inside it).' (Applied effector)

Not surprisingly, similarities between case and word order are most obvious when the case itself is 'pragmatic', i.e. originates from discourse markers.

A similar point can be made with respect to agreement, which is often seen as an alternative way of marking syntactic relations (cf. Whaley's scale, on which they are complementary dimensions). Some languages indeed provide evidence for this view; for example, in Ik case distinctions are neutralized if the subject is cross-referenced on the verb through agreement (König, Chapter 50). Again the parallel is incomplete, specifically since agreement shows more similarities to the indexing function than to the distinguishing function of case (Siewierska and Bakker, Chapter 19). This is also obvious from the fact that, unlike case, agreement does not show the markedness reversal effects subsumed under Silverstein's generalization. That is, while in a split ergative language ERG case may signal an atypical (e.g. inanimate) agent, agreement rather cross-references typical (e.g. animate) As (Croft 1990). However, in other respects these mechanisms are indeed parallel. Again diachrony is partly responsible for similarities since agreement generally develops from case-marked clitics (see Spencer, Chapter 12, for more discussion).

Like case, agreement may have different functions; some uses are clearly syntactic, and comparable to structural case (for instance, subject agreement in European languages goes hand in hand with nominative case), but other uses are semantic (cf. agreement in Koasati and other active languages), or basically pragmatic (e.g. topic agreement rather than subject agreement) as has been argued for some Bantu languages by Morimoto (2008). Thus, the case-marking system of Koasati represents a mixture of semantic alignment, as manifested by verbal agreement, and accusative alignment, as manifested by case (Kimball 1991; cf. Mithun 1999). In the following example from Koasati the choice of dative agreement is determined semantically (it marks experiencer arguments), while the subject takes the nominative case in accordance with its syntactic function.

Koasati (Kimball 1991: 253–4):

(18) *(Anó-k) ca-hó·p*
(I-NOM) I.DAT-sick
'I feel sick/I am hurt'

There are further parallels between agreement and case. For example, Tlapanec, where case relations are expressed through head-marking of the verb (agreement morphology), shows differential case-marking patterns familiar from other languages with bona fide (dependent-marking) case (see Wichmann, Chapter 56). Further, certain patterns in Tlapanec are reminiscent of global case-marking. Thus if P is less affected it can affect cross-referencing of the A participant, which is

now represented by a special 'pegative' paradigm (see Malchukov and De Swart, Chapter 22, for discussion of similar mismatches in the domain of case). Indeed, one might expect that head-marking would be more prone to global marking compared with dependent-marking. Thus, direct/inverse marking is quite frequently found in verb agreement paradigms (Zúñiga 2006), but in case-marking systems it is rather rare. One example comes from Yukaghir, where ACC is restricted to an inverse configuration (Maslova, Chapter 55). Another example is Ik (König, Chapter 50), where clauses with first/second person subjects display 'case anomaly', ACC being replaced by NOM. In most cases, however, case is a 'local' strategy used for the encoding of certain features (role/prominence) of the host NP, while verb-marking is more suitable for encoding 'global' properties sensitive to characteristics of both arguments. This is expected on an iconicity approach assuming a matching relation between form and meaning (cf. Malchukov 2006a on the role of the Relevance Principle for constraining the encoding of transitivity parameters).

The interrelations between case/adposition-marking, agreement, and word order in different functional domains may be schematically summarized as follows (where the solid lines indicate the central functions of a specific coding strategy and the broken lines peripheral functions):

(19)

	semantic	syntactic	pragmatic
case/adposition	———	———	- - - - - - - - -
Agreement	- - - - - - - - -	———	- - - - - - - - -
word order		- - - - - - - - -	———

This does not exclude the possibility that certain strategies might be used in extended functions, but we would predict that in this case the strategy would be used in its central functions as well. Indeed, marking semantic functions is a universal property of cases and adpositions, even though some languages can extend their use to pragmatic functions as well (e.g. marked nominative languages). Conversely, word order is universally used for encoding discourse-pragmatic notions, but although some languages (especially SVO languages) may extend its use for identifying subjects and objects, word order is hardly ever used to encode semantic roles.[10] Agreement has a syntactic and/or pragmatic function in cases when one argument is cross-referenced, but may perform semantic functions as well, in particular, in languages where verbs show multiple agreement paradigms (cf. the Koasati example above).

[10] The closest case would be restrictions on word order in languages such as Movima (Haude 2006), where animates consistently precede inanimates in a clause; yet here the relevant factor is prominence rather than semantic role per se.

45.5 Conclusions

In this chapter we have provided a survey of typological variation in case systems. As is clear from this presentation, case systems can vary in terms of their inventory (see section 45.1), in terms of the mapping between morphological and syntactic case (45.2), in terms of the functional coherence in the case paradigms (45.3), as well as in the interaction between case and alternative strategies (45.4). It was shown that variation in inventories of case systems reflects the order in which cases grammaticalize. Further, case paradigms are more often than not functionally heterogeneous. This is presumably due to diachronic factors, and to the fact that core cases may be recruited from discourse markers. Another source of cross-linguistic variation concerns the mapping between syntactic and morphological case: it has been shown that a strict separation of these notions is necessary to resolve some controversial issues in alignment splits. Finally, languages differ in how case-marking of arguments relates to alternative strategies such as agreement and word order: both partial functional overlap and partial complementarity can be explained with reference to the primary functions of individual strategies (see also Siewierska and Bakker, Chapter 19).

Acknowledgements

We are grateful to Barry Blake and Denis Creissels for helpful comments on an earlier version of this paper. Andrej Malchukov gratefully acknowledges the financial support of the Netherlands Organization for Scientific Research (NWO) for the research reported here (grant no. 220-70-003 'Case cross-linguistically').

CHAPTER 46

CASE MARKING IN DAGHESTANIAN

LIMITS OF ELABORATION*

MICHAEL DANIEL
DMITRY GANENKOV

46.1 Introduction

This paper provides an overview of case systems in Daghestanian languages. To facilitate the presentation of the data, we will start with a brief introduction to the classification of the Nakh-Daghestanian, of which Daghestanian is a regional subset (rather than a genetic subgroup). The family includes several branches – Nakh, Tsezic, Avar-Andic, and Lezgic. Dargwa and Lak are traditionally considered either family-level isolates or forming a separate Lak-Dargwa branch of Nakh-Daghestanian.

Nakh includes Chechen and Ingush, both major written languages spoken in the Republics of Chechnya and Ingushetia, respectively; and Bats, alias Tsova-Tush, an unwritten minority language spoken in Georgia. Data from these languages will not be used below.

* We dedicate this paper to our teacher Aleksandr E. Kibrik, who initiated us into the lore of Daghestanian languages, and on whose various publications this paper is built to a great extent.

978–0–19–920647–6 47-Spencer-ch46 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 669 of 685 May 16, 2008 19:23

Tsezic is divided into East Tsezic, including Hunzib and Bezhta, and West Tsezic, including Tsez, Ginukh, and Khvarshi, all unwritten. Tsezic languages are spoken in the northwest Daghestan near the Chechnya border. Avar-Andic includes Avar, the most widespread language of Daghestan with a number of dialects, not always mutually understandable, spoken in Central Daghestan; and Andic languages, including Akhvakh, Andi, Botlikh, Chamalal, Godoberi, Karata, Bagvalal, and Tindi, all minority unwritten languages, spoken in the northwest of Daghestan near the Chechnya border, to the north of the Tsezic speaking area. Lak is a major language spoken in Central Daghestan; to the southeast of Avar. Various dialects of another major language, Dargwa, sometimes considered to be separate languages that constitute a group on their own, are spoken in the eastern Daghestan. Lezgic languages are spoken in the south Daghestan, near the Azerbaijani border, including the most well-populated representative of the group, Lezgian, and then Agul, Tabassaran, Rutul, Tsakhur, and, in Azerbaijan itself, Kryz, Budukh, Udi, and Khinalug. In Azerbaijan, there are also important Lezgian, Rutul, Tsakhur (as well as minor Avar and Akhvakh) communities. Khinalug, traditionally classified as Lezgic, is considerably different from the rest of the group, and is sometimes considered to be a family-level isolate. Another Lezgic language, Archi, spoken well to the north from the Lezgic area, in central Daghestan, mostly surrounded by Avar- and Lak-speaking villages, also manifests important structural differences from other languages of the group.

For the sake of brevity, we do not insert reference to a grammar of a language each time we quote data. If there is no source indicated, the data comes from one of our 'main sources', i.e. basic grammars of the language. These main sources include:

Tsezic: Bezhta (Kibrik and Testelets 2004); Hunzib (van den Berg 1995); Ginukh (Lomtadze 1963).

Avar: Bokarev (1949a); Alekseev and Ataev (1998).

Andic: Godoberi (Kibrik 1996), Chamalal (Bokarev 1949b), Bagvalal (Kibrik 2001), Akhvakh (Magomedbekova 1967); Karata (Magomedbekova 1971).

Lak: Zhirkov (1955), Murkelinskij (1971).

Dargwa: Abdullaev (1954); Magometov (1963, 1982); van den Berg (2001); Mutalov (2002); Sumbatova and Mutalov (2003).

Lezgic: Lezgian (Haspelmath 1993a); Tabassaran (Magometov 1965); Agul (Magometov 1970; Merdanova 2004); Tsakhur (Talibov 1979; Ibragimov 1990; Kibrik 1999); Rutul (Ibragimov 1978; Makhmudova 2001); Kryz (Authier, n.d.); Archi (Kibrik et al 1977); Udi (Gukasian 1974; Schulze, n.d.).

Khinalug: Kibrik et al. (1972).

Other references are made explicitly in the text. The data that comes from authors' own fieldnotes is marked (f.n.). A more comprehensive list of references on Daghestanian languages may be found in van den Berg (2005).

46.2 Overview

Daghestanian languages are world-famous for the richness of their nominal paradigm. A recently published paper by Aleksandr Kibrik is even entitled 'Nominal Inflection Galore . . . ' (Kibrik 2003a). Indeed, the richest systems, such as those of some Tsezic languages, count up to some seven dozens of forms (within each number value). This is due mostly to the impact of the spatial forms that express such categories as localization (*under the bed*) and movement (*onto the wall*) (see Kibrik 1970, Comrie and Polinsky 1998). Daghestanian case systems are very consistently ergative. They also show a tendency towards high semanticization of case forms. One case form typically covers a set of semantically similar role-marking functions (valency rearranging processes like passives, applicatives, and antipassives are very marginal); see Kibrik (2003a, b). They tend to mark Experiencer separately from both Agent and Patient to a much greater degree than e.g. standard average european, and some of them even use a dedicated case marker, affective. Similarly, most languages have a dedicated construction or, rarely, a dedicated case marker to mark involuntary Agent.

46.3 Paradigm structure

There is a common type of paradigm visible in most Nakh-Daghestanian languages, which manifests several general principles. Declension mostly follows a two-stem pattern: all cases except nominative[1] are derived from a common stem called oblique, while the nominative is derived from a direct stem and is most often formally identical to it (thus being zero-marked). The oblique stem is derived from the direct stem by adding various morphemes called oblique stem markers. The choice of the oblique stem marker is considerably lexicalized, although, sometimes,

[1] Note that, following A. E. Kibrik, we use the term nominative for the case marking of the S/P role in ergative languages (rather than absolutive).

Table 46.1. Lezgian: oblique stem formation and diversity

	paradigm structure for *balk'an* 'horse'		oblique stem markers	
	Singular	Plural	Singular	Plural
Nominative	balk'an horse	balk'an-ar horse-PL		
Oblique stem	balk'an-di- horse-OBL-	balk'an-ar-i- horse-PL-OBL	-di-, -ci-, -c'i-, -či-, -č'i-, -ži-, -ra-, -re-, -ni-, -i-, -u-, -ü-, -a-, -e-	-i-
Genitive	balk'an-di-n horse-OBL-GEN	balk'an-ar-i-n horse-PL-OBL-GEN		
Dative	balk'an-di-z horse-OBL-DAT	balk'an-ar-i-z horse-PL-OBL-DAT		

some phonotactic (syllable count, direct stem auslaut) and, more rarely, semantic correlations may be observed; most languages also have a default oblique stem marker, used whenever there are no phonotactic or semantic preferences. The derivation is similar in both singular and plural; however, the number of oblique stem markers is usually much lower in the plural.

In some languages the oblique stem is most often formally identical to one of the oblique cases, ergative (most Lezgic, Avar, some Tsezic, Dargwa) or genitive (Kryz); thus, Lezgian oblique stem *balk'andi-* in Table 46.1 is formally identical to ergative *balk'andi*. There is a great deal of variation in the choice of oblique stem markers; often, nouns have two variants of oblique stem, as Bagvalal *ʕužruq'* 'hedgehog', OBL *ʕužruq'-i-* or *ʕužruq'-u-*. Oblique stem formation may also be irregular, especially in pronouns; cf. Archi 'I' which has nominative *zon*, ergative *zari*, genitive *-is/-as-* and dative *-ez* (the latter two forms have class agreement prefix), and other cases formed on the oblique stem *za-*.

Some languages may show correlation between the choice of the oblique stem markers and the agreement class of the noun (e.g. masculine and feminine oblique stem markers in Avar-Andic and Archi) or other semantic properties of the noun (e.g. oblique stem marker *-ala-/-ela-* in Agul, characteristic of utensils and other instruments, although this correlation is limited to monosyllabic stems). These patterns, however, never represent a major declension pattern in any of the languages.

At least some declension types in a language may deviate from the Daghestanian prototype. In some languages and dialects one-stem declension is dominant (Khinalug, Udi, Gigatli Chamalal) or well represented (Akhvakh, Godoberi, Tsez, Khvarshi, Tsakhur), with all cases produced directly from one stem (usually identical to the nominative) (Kibrik and Kodzasov 1990). On the other hand, Dargwa oblique cases derive from an oblique stem, but genitive is formed independently from the direct stem; in one of the Tabassaran declension types the oblique

Table 46.2. Deviations from the prototype

Godoberi		Dargwa		Tabasaran	
one stem declension		genitive derived from direct stem		ergative derived from direct stem	
Nominative (direct stem)	hanqu house	Nominative (direct stem)	ganzi stairs	Nominative (direct stem)	č'al language
Oblique stem	hanqu- house	Oblique stem	ganzi-li stairs-OBL	Oblique stem	č'al-na- language-OBL-
Genitive	hanqu-li house-GEN	Genitive	ganzi-la stairs-GEN	Ergative	č'al-nu language-ERG
Dative	hanqu-łi house-DAT	Dative	ganzi-li-s stairs-OBL-DAT	Dative	č'al-na-z language-OBL-DAT

stem is marked by *-na*, *-ra* while ergative is formed from the direct stem by adding *-nu*, *-ru*, *-lu*.

Note that these cases could have derived from the prototype by some diachronic process, such as contraction in the case of Dargwa. Another mode of deviation is attested in Bagvalal, where the vocative suffix is attached to the nominal stem. In case this formation is considered to be a case form, it must be admitted to be produced from the direct stem along with the nominative (the same is true of the Bagvalal 'generalized locative' *-łа*).

A rare typological feature of the Daghestanian languages is the presence of an agreement position in some case markers, including genitives in human noun declension in Bagvalal (controlled by the class of the head) or affective in Andi and Tukita Karata and some locative forms in Dargwa and Lak (all controlled by the class of the nominative core argument); cf. Dargwa 'under the chair' *uta-li-u-w* (chair-OBL-SUB-M).

46.4 Inventory of non-local cases

A case paradigm consists of non-local (alias grammatical) cases and spatial forms. All Daghestanian languages distinguish at least three non-local cases – nominative, ergative, and dative (except some southern Dargwa dialects that lack dative). The dative, however, may be more or less widely used in locative (lative) functions. This is an interesting parameter of variation across the languages of the family, from

978–0–19–920647–6 47-Spencer-ch46 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 673 of 685 May 16, 2008 19:23

Bagvalal (dative with spatial postpositions) to Archi (dative with spatial postpositions and verbs of contact, like 'hit') to Khvarshi (where the lative function is at least as frequent as Recipient / Beneficiary marking; according to Zaira Khalilova, p.c., for this language it is more appropriate to call it lative).

Most languages also have genitive (disputable exception is Tsakhur), and some even have two different genitives, see below on their syntactic distribution. Lak genitive is, however, homophonous with the ergative.

In some languages, personal pronouns do not distinguish between ergative and nominative: this is true of all personal pronouns 'I', 'you.SG', 'we', 'you.PL' (most Tsezic and Lezgic, some Andic); only singular pronouns 'I' and 'you.SG' (in Tsez, Andi, and a dialect of Karata); only plural pronouns (Khinalug, Gakvari Chamalal); and only second person pronouns 'you.SG' and 'you.PL' in Archi. Avar, Lezgian, Lak, and Dargwa lack this syncretism (Magomedova 1979, Kibrik and Kodzasov 1990).

Only few languages limit themselves to the basic set of the four grammatical cases. However, presence of further cases is a matter of variation between languages and branches. Lezgic languages often introduce comitative, absent from all other branches except Dargwa and, according to Creissels (p.c.), Akhvakh. All Tsezic languages have instrumental; outside Tsezic, instrumental is only reported in Kryz (where it is probably derived from dative) and Dargwa (where it is poorly attested). All Andic languages feature affective, a dedicated case marker that codes Experiencer with some verbs (other experiential verbs use dative). In Akhvakh, the case marker cognate to the affective of other Andic languages is, in addition to Experiencer, also used to mark Recipient (competing with lative in this function). Outside Andic, the affective is only found in Tsakhur. Dedicated comparative, the case of the benchmark of comparison ('I am higher *than you*') is found in e.g. Tsez, Hunzib, Rutul, and Archi.

Further cases are introduced individually, for example substitutive 'instead of' in Bagvalal and Hunzib; involuntary Agent case in Lak and Bagvalal; themative 'about, on the subject of, referring to' in literary Dargwa; causal 'because of' in Akhvakh (Creissels p.c.) and Hunzib. Archi is the absolute champion in extending its non-local case paradigm by including comitative, comparative ('(smaller) than the horse' *niʔši-xur* horse.OBL-CMPR), substitutive ('instead of you' *wa-ʟ'ǝna* you.sg.OBL-SUBST), causal ('because of the booze' *c'at'i-li-ši* drink-OBL-CAUSAL), elective ('(one) of these seven girls' *wiʟ-aru laha-qʔiš* seven-2 child.OBL-ELECT), and equative ('in the way of, similarly to the poor' *misgin-ni-qʔdi* poor-OBL-EQUAT).

Note that some peripheral non-local forms are clearly connected to locatives formally (e.g. Archi elective *-qʔiš*, which is diachronically related, in some texts even identical, to Inter marker *-qʔ* plus elative marker *-š*, or Bagvalal unintentional Agent marker *-č'ali*, which is probably related to Cont *-č'*) or have residual spatial usages

978–0–19–920647–6 47-Spencer-ch46 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 674 of 685 May 16, 2008 19:23

(Bagvalal affective *-ba* is required by some spatial adverbs and has a lative value with some place names; while Archi comparative *-χur* is peripherally used to designate spatial adjacency).

46.5 Inventory of spatial forms

Unlike many languages of the world where they are mostly expressed by means of adpositions, many spatial meanings in Nakh-Daghestanian are conveyed by bound morphemes and form a subsystem of nominal inflection, fairly rich in most of the languages (Kibrik 1970 and 2003a, Comrie and Polinsky 1998). Paradigmatically, these subsystems are clearly delimited from non-local cases. A spatial form typically includes two separately coded categories, localization and orientation. Localization defines a certain spatial domain with respect to a landmark – see the examples in Table 46.3.

Here the house is the landmark, which defines three spatial domains 'behind', 'in front of', and 'inside', coded by three localization markers, *-q* (glossed Post), *-h* (Ante) and *-ʔ* (In), respectively.

Orientation conveys the notion of movement, indicating direction of the motion of an object with respect to the spatial domain specified by the localization marker. Central values of the orientation category include motion from the domain (elative), motion to the domain (lative), less often motion towards the domain (allative), motion through the domain (translative) and motion until reaching the domain (terminative). See Table 46.4.

The orientation marker thus necessarily requires the presence of a localization marker, while the opposite is not true – absence of an orientation marker (or, under an alternative interpretation, zero-marked orientation) indicates absence of movement (essive). One rare exception is Dargwa, where essive is more marked than lative, differing from the latter by the presence of a class agreement marker: 'onto the chair' *uta-li-či* (chair-OBL-SUPERLATIVE), but 'on the chair' *uta-li-či-b* (chair-OBL-SUPER-N).

Table 46.3. Agul: three localizations

χul-a-q	χul-a-h	χul-a-ʔ
house-OBL-POST	house-OBL-ANTE	house-OBL-IN
behind the house	in front of the house	in the house

Table 46.4. Agul: three localizations x three orientations

χul-a-q house-OBL-POST behind the house	χul-a-q-t̄i house-OBL-POST-LAT (to) behind the house	χul-a-q-as house-OBL-POST-ELA from behind the house
χul-a-h house-OBL-ANTE in front of the house	χul-a-h-t̄i house-OBL-ANTE-LAT (to) in front of the house	χul-a-h-as house-OBL-ANTE-ELA from in front of the house

The number of localizations in Nakh-Daghestanian varies from four (e.g. Tsakhur) to eight (Agul) or nine (Bezhta and Tsez), typically including meanings such as 'inside', 'on (the surface)', 'behind', 'near', and 'under', more rarely 'in front of' (only Agul and some southern Dargwa dialects). Languages tend to further specify some of these relations by splitting them into two localization categories – especially 'inside' and 'on', sometimes also 'near' (Kibrik 1970).

Thus, the 'inside'-relation often splits into two localizations, In vs. Inter. Some languages use the distinction simply to classify types of landmarks, as Archi, Avar, and most Andic languages, where In conveys the meaning 'inside a hollow object' (containers such as house or mug), while Inter means 'inside a mass object' (such as flour or water). Inter-forms of the names of containers and In-forms of the names of mass objects are ungrammatical in these languages. In Agul and Lezgian the distinction is similar but optional – In may be used with names of both containers and mass objects (when the structure of the latter is irrelevant), while Inter is limited to mass objects and is only used to emphasize the mass character of the landmark. Similarly to Agul and Lezgian, Tabassaran allows both containers and mass objects to combine with In. Additionally, it extends Inter to containers, using the choice between the two forms (In vs. Inter) to convey the distinction between 'loose' (default) and 'close' containment. Close containment means that the object occupies the whole of the inner space of the container ('the wardrobe is full with clothes') or hardly enters in it ('the child hid in / squeezed himself into a box') or is fixed in it ('the glass is inserted into the window frame'). In Akusha and some other dialects of Dargwa Inter is impossible with the names of containers; the use of In with the names of mass objects indicates that the landmark includes an object as its element, such as 'The sand contains stones; the sand is stony', while Inter is reserved for simple physical location ('the stones are in the sand').

Another frequently occurring split is between two types of 'on'-relation: localization Super vs. localization Cont. Traditionally, this distinction is considered to convey the opposition between location on horizontal vs. vertical surfaces. In fact the portrait of this distinction is much more complicated. To give some examples, in Agul, Super is used for location on a supporting surface (such as a book lying

Table 46.5. Distribution and functions of the two 'inside' localizations: In vs. Inter

	Avar		Dargwa		Tabassaran	
	mass object	container	mass object	container	mass object	container
In		default	element of a composite structure	default	landmark structure irrelevant	default
Inter	default		default		default	close containment

on the table), animate objects autonomously keeping on the landmark (a fly sitting on the wall/on the ceiling), as well as objects being part of the surface (a scratch on the mirror) or natural extensions of the landmark (as a leaf on the branch). Cont, on the other hand, is used for objects attached to, rather than being part of, the landmark, such as a painting hanging on the wall, or being a characteristic of the landmark (such as a beard on the cheeks – bearded cheeks, or for meanings like 'there are raindrops on the window, the windowpane is covered with raindrops'). In Tsez, Super is limited to supporting surfaces and surfaces including the object as its element. All other 'on'-relations are expressed by Cont.

For a more detailed discussion of 'split localizations' see Ganenkov (2005).

A further complication is that there is a number of non-spatial or not straightforwardly spatial meanings formally integrated into the spatial sub-paradigm as additional localizations. Some languages have a dedicated human locative (HumLoc) localization which conveys the meaning of being located in someone's personal space (house etc.), attested in Bagvalal and Archi. In Tindi and Bezhta (and, probably, Karata), there is a dedicated possessive localization marker, used in predicative possessive constructions (possessive essive) and in 'take-away-from' constructions (possessive elative).

Tladal Bezhta (Ganenkov, f.n.)

(1) *dī-qa* *oqro* *gel*
I.OBL-POSS money(NOM) COP
'I have money (on me).'

Some languages contribute further categories to the typological profile of the Daghestanian spatial morphology, already extremely rich. Thus, in Bezhta and Hunzib, there is a marker described as approximate location, while Kaitag Dargwa has morphologized the expression of vertical and personal deixis; cf. *qalžirk'en* / *qalžiχen* 'up/down the roof'; *qalžirten* / *qalžiržen* 'thither / hither by the roof'.

As was already mentioned before, the richness of the spatial paradigm in Daghestanian is mostly due to the combinatorial regularity of a limited set of markers. However, this regularity must not be considered to be absolute. It is relatively common not to distinguish between lative and essive with some localizations, while distinguishing them in the others. Thus, Andi distinguishes lative and essive in Apud, In, and Super, but doesn't in Cont, Sub, Inter, and Ad (Magomedova 1979). Gigatli Chamalal distinguishes between these two orientations everywhere except Cont, while in Rutul, on the contrary, only Cont makes the distinction. In Agul and Lezgian, lative does not combine with 'In', and dative or a special combination of dative plus lative is used instead (while normally dative does not combine with orientation markers). Archi is exceptional in that it has no essive or translative for the Cont localization. In Karata, there is a human lative 'to someone's place' in *-χar*, in which the lative marker *-ar* may be isolated; however, the putative *-χ-* localization does not combine with any other orientation marker, synchronically.

The distinction between grammatical cases and spatial forms is by no means purely functional, in the sense that some of the latter have widespread syntactic usages. For example, the Stimulus of 'fear, be afraid of' is coded by Sub-Elative in Archi, Apud-Elative in Tabassaran, Post-Elative in Lezgian, Ante-Elative in Agul, Super-Essive in Hunzib (Testelets 1980), Cont-Essive in Bagvalal, Cont-Elative in Andi and Godoberi, Super-Elative in Lak and Rutul (Ganenkov, f.n.). For further discussion see Kibrik (2003a).

46.6 Place names and natural locations

Speaking of space in Daghestanian, it is hard not to mention place names. Local place names (such as names of neighbouring villages) often have a reduced case paradigm, being limited to spatial forms. Thus, in e.g. Bagvalal, there are place names that, of all non-local cases, only have genitive; essive serves as their citation form. In an argument position, instead of e.g. nominative, a periphrastic construction with some kind of hyperonym is used, such as 'village' or 'place', as in the following example.

Bagvalal (Daniel 2002)

(2) *di-č'* *k'^{w}an-ɬ* *han* *raq'^{w}a-ɬi* *ek'^{w}a*
I.OBL-CONT Kvanada-GEN village(NOM) heart-INTER COP
'I remember Kvanada' (lit. 'The village of Kvanada is in my heart').

Even inside the spatial sub-paradigm, place names may behave differently from regular nominals. Typically, they either do not have a localization marker at all

or combine but with a single localization marker, lexically defined (and probably historically motivated), as in Avar-Andic or Dargwa (Avar 'in Khunzakh' *χunza-q* Khunzakh-APUD, but 'in Holotl' *holo-l* Holotl-SUB), or with a default localization marker, as in Lezgian and Agul (Agul 'in Tpig' *tipʷaʁ-a-ʔ* Tpig-OBL-IN). They distinguish orientations only (essive vs. lative vs. elative) and, in this respect, behave very much like the spatial adverbs (which typically use the same set of orientation markers as the nominals but have no localizations). Further, in e.g. Agul, lative of place names is morphologically unusual, being formed immediately from the direct stem, with no oblique stem or localization marker ('to Tpig' *tipʷaʁ-di* Tpig-LAT), unlike essive and elative, which have both. All these phenomena clearly reflect the intimate relation between place names and spatial semantics.

Another class that shows spatial irregularities are names of locations, such as 'landscape elements' ('field', 'village', 'cave'), buildings ('house', 'cowshed'), bodily locations ('hand', 'skirt hem', 'armpit'). These nominals, naturally occurring in locative contexts, tend to preserve older spatial morphology longer than other nominals and thus become in some way irregular, for example possessing a separate 'locative' stem (identical to the former oblique stem, now lost), which is different from both the direct and oblique stem (Archi), or a specific localization marker (Bagvalal).

46.7 Black sheep of the family

What was described earlier may be called the Daghestanian prototype. However, there are some Daghestanian languages that lack such typically Daghestanian features as the spatial subsystem in the nominal inflection, or almost do so. South Avar dialects have lost the elaborate locative morphology. Khinalug and Budukh also have but a very reduced version of the system. Vartashen Udi does not show any vestiges of the older spatial sub-paradigm, and Nidzh Udi evidence is not much stronger. Observe the Nidzh Udi case suffixes (Table 46.6).

Table 46.6. Nidzh Udi: inventory of cases

nominative	(zero)	ablative/comitative	*-χun*
ergative	*-en/-in*	adessive	*-sta*
benefactive	*-jnak*	allative	*-č*
genitive	*-e/-a/-in*	superessive	*-l*
dative	*-a*	superelative	*-lχun*

978–0–19–920647–6 47-Spencer-ch46 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 679 of 685 May 16, 2008 19:23

Note that Super-Elative is formally identical to super(essive) plus ablative, which is reminiscent of Daghestanian locatives. However, all the locative forms except ablative are not in common usage; and Super-Essive and Super-Elative are extremely rare.

46.8 Core arguments

In terms of case marking, all Nakh-Daghestanian languages manifest ergative alignment. Verbs fall into two major groups, intransitive and transitive, of which transitives take an argument in ergative (Agent) and another argument in nominative (Patient), while intransitives take a nominative argument but do not take an ergative argument. Most Nakh-Daghestanian languages also have labile verbs that have both patterns, transitive and intransitive. Some intransitives may have additional arguments, typically in one of the spatial cases. Cf. Lezgian examples:

Lezgian (Haspelmath 1993a)

(3) *am* *gamiš-di* *ja-na*
that(NOM) buffalo-OBL(ERG) gore-PAST
'A buffalo gored him.'

(4) *χalq'* *wič-i-n* *ǭuwat-di-q* *inanmiš* *tir*
people(NOM) REFL-OBL-GEN power-OBL-POST believe become.PAST
'The people believes in its own force(s).'

Different from the latter are experiential verbs, which include 'see', 'hear', 'know', 'love, want', sometimes also 'find', 'forget', and consistently use dative or, for some verbs in some Andic languages and Tsakhur, a dedicated, affective case marking for the Experiencer; the Stimulus is expressed by nominative or a complement clause (cf. discussion in Kibrik 2003a, Ganenkov 2006).

Bagvalal (Kibrik 2001)

(5) *he<b>ō* *du-ha* *q'oča-m-o* *ek'^{w}a ?*
what<N> you.SG.OBL-DAT want-N-CVB COP
'What do you want?'

(6) *basqan* *ʕumar-i-ba* *uha-m-o* *ek'^{w}a* *han*
Baskan Omar-OBL-AFF be.able-N-CVB COP village(NOM)
b-uh-ā
N-gather-POT.INF
'Baskan Omar managed to capture the village.'

Many languages have a special intransitive construction introducing an Agent-like participant with limited agentive properties (cf. Ganenkov et al. 2006; Kittilä 2005b), typically coded by one of the spatial cases or, rarely, by a dedicated case form, as in Lak and Bagvalal.

Archi (Daniel, f.n.)

(7) *za-ra-š* *č'ut* *a<b>qI-u*
I.OBL-CONT-ELA jar(NOM) <III>break-PFV
'I broke a jar (occasionally).'

Bagvalal (Kibrik 2001)

(8) *di-č'ali* *o-w* *w-ič'a*
I.OBL-INVOLUNTARYAGENT that-I(NOM) I-die
'He died because of me, I killed him unintentionally.'

Ditransitive constructions follow the 'indirect object' pattern in all languages, coding the Theme (the object transferred) by nominative and the Recipient by dative. Some Dargwa dialects have no dative and use a spatial form instead, as Super-Lative in Icari Dargwa.

Icari Dargwa (Sumbatova and Mutalov 2003)

(9) *q'^{w}aIl-li-j* *mura* *saka<b>iχ̄-a*
cow-OBL-SUP(LAT) hay(NOM) <N>put.PFV-IMP
'Give the hay to the cow.'

(10) *du-l* *cin-na* *qal* *c'a-l* *b-iǩ̄-ub*
I-ERG REFL-GEN house(NOM) fire-ERG N-burn:PFV-PRET
admi-li-j *azir* *q̄uru* *b-īč'-ib-da*
person-OBL-SUPLAT thousand ruble(NOM) N-give:PFV-PRET-1SG
'I gave one thousand rubles to the man whose house had burnt down.'

Most languages further distinguish between 'give'-situations that involve transfer of possession ('give, donate') and those that do not involve it ('lend, give back'). See examples from Archi.

Archi (Kibrik 1977)

(11) *zari* *wa-s* *aL'* *Lo*
I.ERG you.SG.OBL-DAT meat(NOM) 4.give.PFV
'I gave you the meat (for good).'

(12) *za-ra-k* *jamu-t* *he'ǝna* *aǩ̄u-s* *oq'i*
I.OBL-CONT-LAT this-IV thing(NOM) 4.see-INF 4.give(IMP)
'Give me this thing so that I (can) look at it.'

(13) *zari* *to-w-mu-ra-k* *q'onq'* *baq'Ia-s* *aw*
I.ERG that-1-OBL-CONT-LAT book(NOM) go.back-INF 4.do.PFV
'I returned him the book.'

Causative constructions. Nakh-Daghestanian languages are rich in various causativization patterns. The Causer is consistently marked by ergative, and the Patient of the originally transitive verb is coded by nominative. The Causee marking depends on the (in)transitivity of the original, non-causative verb: a Causee originating from the only argument of an intransitive (intransitive Causee) is coded by nominative, while a Causee that used to be a transitive Agent (transitive Causee) is coded by one of the spatial cases. See the Agul examples.

Agul (Daniel, et al. to appear)

(14) *baw-a* *šünük̄* *ʀarx-a-s* *q'-u-ne*
mother-OBL(ERG) child(NOM) sleep-IPFV-INF do-PFV-PF
'Mother made the child sleep.'

(15) *malla-ji* *gada-ji-w* *q'urʔan* *ruχ-a-s*
priest-OBL(ERG) boy-OBL-APUD Koran(NOM) read-IPFV-INF
q'-a-a
do-IPFV-PRES
'The priest makes his son read the Koran.'

Interestingly, Agul has other options of coding the Causee. Intransitive Causee may be coded by Apud, emphasizing the decrease in his or her control over the caused situation, while transitive Causee may be coded by an ergative, marking the increase in this control. The latter marking results in the presence of two ergatives, which makes the morphosyntactic status of Agul 'do'-causatives disputable.

Agul (Daniel, et al. to appear)

(16) *dad-a* *uči-n* *uq̄ub-ar-i-l-di* *gada-ji-w*
father-ERG REFL-GEN beating-PL-OBL-SUP-LAT son-OBL-APUD
χula-as *hiš-a-s* *q'-u-ne*
house-IN.ELA flee-IPFV-INF do-PFV-PF
'Father's beating made (his) son run away from home.'
(lit. 'by his beating father made son run away from home')

(17) a. *gi* *šünük̄-ar-i* *wak̄-a-n* *jak̄* *ʕut'-a-s*
that(ERG) child-PL-ERG pig-OBL-GEN meat(NOM) eat-IPFV-INF
q'-u-ne.
do-PFV-PF
'He let children eat pork.'
(e.g. he forgot that they are Muslims, or neglected the dietary restrictions)

b. *gi* *šünük̄-ar-i-w* *wak̄-a-n* *jak̄* *ʕut'-a-s*
that(ERG) child-PL-OBL-APUD pig-OBL-GEN meat(NOM) eat-IPFV-INF
q'-u-ne.
do-PFV-PF
'He made children eat pork.'
(e.g. although, being Muslims, they didn't want to)

Nominative requirement. Nakh-Daghestanian clauses typically require one and only one nominative NP to be present, although this fact is dimmed by a more or less extensive pro-drop. However, there is a number of recurrent exceptions. Meteorological predicates may have no arguments at all (in Bagvalal, an Arabic loan *dunijal* 'universe' may optionally be inserted into the nominative slot; cf. (18)). With verbs with 'default objects', the Patient-nominative may be omitted because the object is predictable from the verbal semantics, while any non-default object must be expressed; cf. (19). With 'hitting'-verbs it is often the hitting object (rather than object or person hit) that is conceptualized as Patient-nominative; this instrument-like nominative may be omissible; cf. (20).

Bagvalal (Kibrik 2001)

(18) *[dunijal]* *ʁori*
[world(NOM)] thunder.strike
'Thunder struck.'

(19) a. *den* *turi*
I.ERG spit.PAST
'I spat (saliva) / I spat out a fruit stone'

b. *den* *raL'* *turi*
I.ERG stone(NOM) spit.PAST

Agul (Ganenkov, f.n.)

(20) *fas* *jarħ-a-j-e* *ramiz-a* *ħajwan-di-s*?
why hit-IPFV-CVB-COP Ramiz-ERG horse-OBL-DAT
'Why is Ramiz beating the horse?'

Finally, a nominal stem may be closely related to the verb in terms of the participant structure. These stems do not form fully fledged NPs in the sense that they may not have adnominal dependents (nouns or adjectives). Still, they may behave differently in being fully integrated with the verb, morphosyntactically, and thus freeing the nominative slot for another noun, as in (21), or keeping the nominative slot for itself without letting any real Patient or Patient-aligned argument occupy this position, as in (22).

Agul (Daniel, et al. to appear)

(21) *wa-s* *haraj-ar* *un* *x-u-ne-wa*?
you.SG.OBL-DAT shout-PL(NOM) sound become-PFV-PF-Q
'Have you heard the yelling?'

(22) *wa-s* *bawa* *un* *q'-a-a*?
you.SG.OBL-DAT mother(ERG) sound(NOM) make-IPFV-PRS
'Mother calls you' (lit. 'mother to you sound makes')

Another type of context where two nominatives are present are analytic forms of transitive verbs, where both Agent and Patient may be coded by nominative,

978–0–19–920647–6 47-Spencer-ch46 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 683 of 685 May 16, 2008 19:23

resulting in the so-called binominative construction, fairly widespread in Daghestanian. Their function varies across languages, and their morphosyntactic status is subject to debate.

46.9 Adnominal and predicative possession

Genitive is a quasi-universal case form in Daghestanian.

There is one language where the existence of the genitive case is disputed. In Tsakhur, the 'genitive' marker is attached to various lexical categories, including plain nouns, nominal case forms, adjectives, and finite verbal forms (forming relative clauses) and is thus qualified as a transcategorial attributivizer. On the other hand, with nouns, this attributivizer is added to oblique stems, which is a property specific only to the members of the case paradigm.

Distinction between alienability and inalienability is only attested in Budukh (Authier, p.c.) and Khinalug.

Khinalug (Kibrik et al. 1972)

(23) *gad-i* *kalla* vs. *gad-e* *c'^{w}a*
boy-GEN.INAL head boy-GEN.ALIEN house
'boy's head' vs. 'boy's house'

In some Tsezic languages, e.g. Bezhta, there are two distinct genitives that are distributed syntactically (Kibrik 1995). One is used with nominative heads, the other with heads in any other case; Tsakhur has a similar distinction using two different attributive markers depending on the case of the head.

Sometimes, the genitive form may function as an NP head and is further declinable.

Bagvalal (Kibrik 2001)

(24) *a-b* *hob* *in-ɬ̄i-ɬ-da* *waša-šu-b-ɬ̄i-ba*
this-N tomb(NOM) REFL-OBL-GEN-PART boy-OBL-GEN-OBL-AFF
b-aɬ̄i-li-b-o *o-w-la* *raq'^{w}a-lā* *w-ā-w-o*
N-similar-VBLZ-N-CVB that-M-PART heart-SUP.LAT M-come-M-CVB
'Because this tomb was similar **to that of my son**, I remembered him.'

Functionally, genitive in Daghestanian covers a relatively wide range of adnominal meanings, including material ('a mug *of copper*'), elective ('one *of them*'), feature

978–0–19–920647–6 47-Spencer-ch46 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 684 of 685 May 16, 2008 19:23

object ('girl *with blue eyes*, blue-eyed girl'), as well as of course such core meanings as possessive, part–whole, and kin relation.

Possessive predication. Daghestanian languages fall into three groups according to how they express possessive predication formally.

a) Possessor in possessive predication is always expressed by a genitive (e.g. Dargwa and Archi)
b) Possessor may be coded either by a genitive or by a spatial form (typical of Andic and Tsezic)
c) Possessor may be coded by two spatial forms, genitive is only used in adnominal possessive constructions (typical of Lezgic)

In (b) and (c), the contrast between the two markings of the possessor is close to permanent vs. temporary possession or general vs. actual ('I have it on me') possession.

Agul (Ganenkov, f.n.)

(25) *za-w* *kant'* *f-a-a*
I.OBL-APUD knife(NOM) APUD-be-PRS
'I have got a knife (with me).'

(26) *če* *χuj-i-q* *jaq'u* *kurc'ul* *q-a-a*
our.EXCL dog-OBL-POST four cub(NOM) POST-be-PRES
'Our dog has four cubs.'

Bagvalal (Kibrik 2001)

(27) *di-b / di-č'* *tup* *ek'^{w}a*
I.OBL-GEN.1 / I.OBL-CONT gun(NOM) COP
'I have a gun' (I possess a gun) / 'I've got a gun' (somebody else's gun, I've got it with me).

Note that in predicative possessive constructions the respective word order of the genitive viz. the Possessee is very free, so the Possessor seems to be syntactically independent.

46.10 Some peripheral roles and functions: highlights

Instrument. Special instrumental case is attested in Tsezic and Dargwa. Other cases that may have instrumental function are comitative (Dargwa, Archi, Tsakhur, Kryz, Budukh), ergative (Avar-Andic, some Lezgic) or a spatial form, as Agul and

978–0–19–920647–6 47-Spencer-ch46 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 685 of 685 May 16, 2008 19:23

Lezgian Super-Lative. Note that many languages have several different ways to mark Instruments, with no clear semantic contrast. In Icari Dargwa, there are even three ways to express Instruments: ergative, comitative, and a dedicated instrumental (the latter in a very limited number of contexts).

Benchmark. Quite a few languages have a specialized form for a standard of comparison, including Tsez, Hunzib, Rutul, and Archi. Other languages express benchmarks by means of spatial forms, such as Super-Elative in most Lezgic, Avar, Chamalal; Cont-Essive in Andi and Bagvalal; Cont-Elative in Godoberi; Super-Essive in Khvarshi; or Ad-Essive in Bezhta.

Addressee. Andic and Tsezic languages and Archi confute the typologically widespread pattern of marking the addressee of speech verbs in the same way as the recipient of 'give'-verbs, coding the former with one of the spatial forms instead, such as Super-Lative in Bagvalal, Ad-Essive in Godoberi, Poss-Essive in Bezhta or Cont-Essive in Tsez.

CHAPTER 47

POOR (TWO-TERM) CASE SYSTEMS

LIMITS OF NEUTRALIZATION

PETER M. ARKADIEV

47.1 Introduction

Two-term (or bicasual) case systems are quite widespread in the world's languages (see section 47.2), but have not received enough attention from linguists. Nevertheless, they constitute an interesting phenomenon whose study may be fruitful from various points of view: bicasual systems show what a 'minimal' case system may be like, which meanings of case markers may go together, how different patterns of argument encoding (accusative, ergative, etc.) may interact under extremely limited expressive possibilities (see sections 47.3–4). Two-term systems are of especial relevance to the diachrony of case systems: very often such systems represent either the last or the first stage of the existence of case in the language, and their study is important for our understanding of how case systems emerge and dissolve (see section 47.5). These are the issues which are going to be briefly discussed in this Chapter.

The following two terminological remarks are in order. First, to avoid confusion and aprioristic labels, the members of a bicasual system will be called Dir(ect)

978–0–19–920647–6 48-Spencer-ch47 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 687 of 699 May 16, 2008 19:0

and Obl(ique). The label Dir is assigned to the case which coincides with the citation form of the noun. Second, this Chapter will be concerned only with those systems where both cases are able to encode semantico-syntactic functions of *verbal* dependents (i.e. *relational* cases); that is, systems where one of the two cases is a Genitive (as in Swedish) or a Vocative (as in literary Bulgarian) will not be discussed here.

47.2 Areal and genetic distribution of two-term case systems

Two-term case systems are attested in almost all major linguistic areas, although their distribution is far from being even. They are sporadically found in Europe, most notably in such already extinct languages as Old French and Old Provençal, but also in Romanian and in the dialects of other Balkan languages, in some Scandinavian dialects, and in modern English pronominals. In Asia bicasual systems are abundant in the Iranian, Dardic, and Nuristani languages, less in the Indo-Aryan languages (see Stilo, Chapter 48, for details), and are also attested in the North-West Caucasian languages (Adyghe and Kabardian). Such systems figure prominently in Africa, where they are found in almost all Berber languages, in the Ethiopian branch of the Semitic family, in many Cushitic languages (all belong to the Afroasiatic phylum), and in the Nilotic languages, which belong to the Nilo-Saharan phylum.

In the New World two-term case systems are not so common, probably due to the overall aversion of these languages towards dependent-marking. Here such systems are attested in the Salish language family, in some Uto-Aztecan languages, in Choktaw (a Muskogean language). In South America two-term case systems are found in a Bolivian isolate Movima, some Chibchan languages, and possibly in the Panoan language Matís, but it is probable that a closer investigation will reveal more such languages in that rather underdocumented region.

In the Pacific area two-term case systems are only sporadically attested, being found in Nias (an Austronesian language of Western Indonesia), in Yimas (a Papuan language of the New Guinea Highlands), in Maung (a Yiwaidjan language of Northern Australia, where case is restricted to independent pronominals), and in Aleut.

Thus, the languages with bicasual systems show great areal and genetic diversity, and it is no surprise that such case systems themselves exhibit considerable cross-linguistic variability. However, commonalities among two-term case systems found in the languages of the world are also quite noteworthy. Both similarities and differences among such systems will be discussed in the next sections.

978–0–19–920647–6 48-Spencer-ch47 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 688 of 699 May 16, 2008 19:0

47.3 Functional properties of two-term case systems

When we look at a case system of a language, there are several questions we must answer in order to characterize it. The principal question concerns the range of meanings the cases are able to express. With respect to two-term case systems this question is rather divided into two:

1. Which semantico-syntactic functions are expressed by cases themselves (and not by adpositions)?
2. How are these functions distributed between Dir and Obl?

Thus, there are two major functional parameters of the typology of two-term case systems:

1. The 'case zone', i.e. the range of functions covered by the cases only, without use of adpositions.
2. The way the functions from the 'case zone' are distributed between the two cases.

Most languages with morphological cases possess also a more or less rich system of adpositions which express various meanings; bicasual systems are no exception, although there are languages with almost no adpositions among them (e.g. Salish and Yimas). We would expect that in a language with only two morphological cases and an array of adpositions the range of functions expressible by cases themselves must be rather limited, since in a two-term case system polysemy may easily lead to ambiguity. What we actually find, however, is that in the overwhelming majority of two-term case systems the 'case zone' is quite broad and usually includes, besides the core roles (S, A, and P; see Haspelmath, Chapter 33, for definitions), also such functions as Addressee/Recipient, Possessor, various locative and other circumstantial relations (cf. Stilo, Chapter 48). A typical system of this kind is found in Old French. Here Dir encodes S of intransitive verbs (see (1), Foulet 1970: 4), A of transitive verbs (see (2), Moignet 1976: 90), and nominal predicate (see (3), Foulet 1970: 8).

(1) *li* *chevalier-s* *s=en* *part.*
the(DIR) knight-DIR.SG REFL=CL departs
'The knight departs from there.'

(2) *il* *vit* *un* *home* *crucefié.*
he(DIR) saw a(OBL.SG) man(OBL.SG) crucified(OBL.SG)
'He saw a crucified man.'

(3) *il* *est* *me-s* *pere.*
he(DIR) is my-DIR.SG father(DIR.SG)
'He is my father.'

978–0–19–920647–6 48-Spencer-ch47 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 689 of 699 May 16, 2008 19:0

Obl, apart from P of transitive verbs (see (2) above), is used to encode the addressee with ditransitive verbs (see (4), Moignet 1976: 91), adnominal possessor (see (5), Foulet 1970: 14), goal of motion (see (6), Moignet 1976: 96), temporal interval (see (7), Moignet 1976: 95), manner of action (see (8), Foulet 1970: 32):

(4) *dites le roi que…*
say:IMP.2PL the(OBL.SG) king(OBL.SG) that
'Tell the king that…'

(5) *la niece le duc*
the niece the(OBL.SG) duke(OBL.SG)
'the niece of the duke'

(6) *droit sentier qui cele part*
direct(OBL.SG) road(OBL.SG) that(DIR.SG) this(OBL.SG) place(OBL.SG)
le menast.
he(OBL.SG) would.lead
'[He could not find] a direct road that would lead him to that place.'

(7) *Erec dormi po cele nuit.*
Eric(DIR.SG) slept little this(OBL.SG) night(OBL.SG)
'Eric slept a little this night.'

(8) *s'=en part le-s gran-z galop-s.*
REFL=CL departs the-OBL.PL great-OBL.PL gallop-OBL.PL
'[And the knight] departs in great gallop.'

Other functions found in two-term case systems include Instrument (see (9) from Squamish, Salish, Kuipers 1967: 169), Location (see (10) from Yimas, Foley 1991: 166), Comitative (see (11) from Movima, Haude 2006: 282).

(9) *na=λič'itas ta=smic t=ta=λač'tn.*
ASP=cut:3SG.A/3SG.P ART=meat OBL=ART=knife
'He cut the meat with a knife.'

(10) *ŋaŋk-ɲan ama-na-irm-n.*
grass-OBL 1SG-DEF-stand-PRES
'I am standing in the grass.'

(11) *kide: da' kaykay jayna n=us alwaja='ne.*
they DUR eat:RDP now OBL=ART spouse=3SG.F
'They are eating now with her husband.'

Functionally rich systems like the Old French one are very common; by contrast, 'narrow' systems, where the case zone is limited just to core cases, or includes only one or two peripheral functions, are rare (cf. the Berber languages, Aleut, and Wakhi, an Iranian language of Pamir). Such a cross-linguistic distribution of

'broad' vs. 'narrow' two-term case systems implies that languages perfectly tolerate extended polysemy of case markers. This is due to the general tendency of highly grammaticalized case markers to encode particular functions only with those nominals which are 'natural' with these meanings (see Aristar 1997). For instance, Obl may be interpreted as 'locative' with names of locations, as 'instrumental' with names of instruments, and as 'dative' or 'comitative' with animate nominals.

Let us turn to the second parameter, the distribution of functions among the two cases. If we first look at the encoding of core grammatical relations, in the languages with bicasual systems we will find all possible kinds of marking. The nominative–accusative marking is most common (e.g. Old French, Uto-Aztecan, Berber, Nilotic, Amharic, Persian, etc.); next comes the neutral encoding (Salish, Yimas, Movima, Aleut). The ergative–absolutive marking is dominant (which does not mean 'unique') in Adyghe and Kabardian, Nias, Matís, and Päri, a Nilotic language, but occurs as an option in Aleut and in many Indo-Iranian languages. The S/A or S/P participant is usually encoded by Dir, but there are notable exceptions, cf. examples (12) and (13) from Kabyle (Berber, nominative–accusative, Chaker 1983: 276, 279), and (14) from Nias (ergative–absolutive, Brown 2001: 94):

(12) *fγ-n y-rgaz-n.*
left-3PL OBL.PL-man-PL
'The men left.'

(13) *y-wt aqšiš-ni w-rgaz-im.*
3SG-hit DIR:boy-this OBL.SG-man-2SG
'Your husband hit this boy.'

(14) *me mofanö ya, la-roro ya niha fefu.*
when left he(OBL) 3SG-follow he(OBL) (DIR)person all
'When he left, everyone followed him.'

In Kabyle, as well as in Nilotic, Cushitic, and Muskogean languages, it is the S/A relation which is marked by Obl, not the P. Similarly, in Nias the A is encoded by Dir, whereas the S/P participant receives morphological marking. The rationale of such systems lies not in the alleged 'unmarkedness' of the S relation (cf. Comrie 1978), but in general markedness principles (cf. Givón 1995, Haspelmath 2006): among the two cases in a minimal system it is the 'default' case used to encode many different functions which remains the unmarked member of the morphological opposition. By contrast, the case whose only function is to encode the 'subject' (S/A) argument, is both functionally and formally marked. The next question is why the 'subject' function is not encoded by the 'default' Direct case in these languages. The answer probably lies in the realm of information structure: it is usually only the non-topicalized and thus functionally marked 'subject' participant which receives Obl encoding in these languages, see the following examples from Tachelhit (Berber, Galand 1964: 34, 40):

(15) *ikrz* *u-rgaz* *igr.*
cultivated OBL-man DIR:field
'[It was] the man [who] cultivated the field.'

(16) *a-rgaz* *ikrz* *igr.*
DIR-man cultivated DIR:field
'[As for] the man [he] cultivated the field.'

Therefore, such 'marked nominative' systems are in fact functionally well motivated (see König, Chapter 35, and König 2006 for a more detailed discussion).

Among the two-term case systems 'split' case marking is very common; in the Indo-Iranian group, where there is both a tense–aspect split in the marking of A, and definiteness/animacy split in the marking of P, up to four constructions (neutral, accusative, ergative, and double-oblique) may co-exist in a single language (cf. Stilo, Chapter 48). A clause-type split is found in the Uto-Aztecan languages, where the S/A participant is marked by Obl in subordinate clauses (see (17) from Yaqui, Lindenfeld 1973: 103).

(17) *na=a* *biča* *ke* *hu-ka* *usi-ta* *čuʔu-ta* *kipwe-ʔu.*
1SG:DIR=3SG see that this-OBL child-OBL dog-OBL have-PRF
'I see that this child has a dog.'

If we turn to the general patterns of the functional organization of two-term case systems, we may find two principal patterns of distribution of meanings between the cases:

1. 'Dividing' systems, where all peripheral functions are attributed to a single case (usually Obl), which may also have a core function;
2. 'Distributing' systems, where both cases have core as well as peripheral functions.

'Dividing' systems are by far the most common (compare Old French above and Table 47.1), while the genuine 'distributing' systems occur only in some languages of the Pamir and Hindukush region, for example in the Nuristani language Kati (see Table 47.2, Èdel'man 1983: 60–1).

Table 47.1. Functions of cases in Old French

Dir	S, A, Pred
Obl	P; Addressee, Possessor, Locative, Goal, Temporal, Manner

Such an uneven distribution of the two types of two-term case systems is probably due to the general tendency of cases to encode 'natural' classes of functions, such as core vs. peripheral or S/A vs. all others. By contrast, in the 'distributing' systems such as that of Kati the only rationale for the 'division of labour' between the cases is their diachronic origin: the Indo-Iranian Dir stems from the collapse

978–0–19–920647–6 48-Spencer-ch47 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 692 of 699 May 16, 2008 19:0

Table 47.2. Functions of cases in Kati

Dir	S, A, P; Goal, Locative
Obl	A in the past tenses, definite P; Addressee, Possessor

of older Nominative and Accusative, while Obl derives from Genitive–Dative. In the 'distributing' systems the two cases retain the functions which belonged to the different cases they originate from. It is noteworthy that the majority of the Indo-Iranian languages must have undergone a functional change and redistributed the functions of cases, so that now their two-term case systems are of a genuinely 'dividing' type.

To summarize, from the point of view of the semantic content of a case system, bicasual systems may be characterized by the following properties:

1. The cases usually cover a broad range of meanings, including both core grammatical relations and peripheral functions (locative, temporal, manner, etc.).
2. The markedness relations between Dir and Obl tend to iconically reflect the functional load of these forms: the case with a greater variety of uses and a non-restricted distribution is usually the morphologically unmarked Dir, even though the 'subject' relation may be encoded by the other case.
3. The distribution of functions between cases more often than not follows the pattern where all peripheral functions are subsumed under one of the cases only.

47.4 Morphological properties of two-term case systems

In the previous section the cross-linguistic trends in the functional organization of two-term case systems were surveyed. Now let us consider the morphological make-up of such systems. Here we find that typologically rare and unusual patterns appear with a frequency greater than average. This concerns both form and position of case exponents attested in bicasual systems. The most frequent type of marker used in such a system is a bound affix, but there are deviations from this prototype in both directions. Thus, in Halkomelem and some other Salish languages, in Amharic, and Persian, Obl is a clitic, cf. example (18) from Amharic where the case marker *-En* behaves as a second-position clitic attaching to the preposed adjective rather than to the head noun (Leslau 1995: 184).

(18) *wəšša=w təlləq=u=n bäqlo näkkäsä.*
dog=ART big=ART=OBL mule bit
'The dog bit the big mule.'

By contrast, in Nias Obl is normally realized as a morphophonological process, that is, 'mutation' of the initial consonant of the stem, cf. Table 47.3 (Brown 2001: 39–40).

Table 47.3. Mutation as case exponence in Nias

	'rice'	'land'	'stick'	'pig'
Dir	*fakhe*	*tanö*	*si'o*	*baßi*
Obl	*vakhe*	*danö*	*zi'o*	*mbaßi*

Morphophonological alternations function as case exponents, usually alongside affixes, also in Old French and Old Provençal, in many Indo-Iranian and Afroasiatic languages. Finally, in Nilotic and Cushitic languages the primary and more often than not the only exponence of case is tone (see Bennett 1974), see the paradigms from Maasai in Table 47.4 (Tucker and Bryan 1966: 459).

Table 47.4. Tone as case exponence in Maasai

	'knife'	'water'	'girl'	'shepherd'	'giraffe'
Dir	*ɛŋkálém*	*ɛŋkáré*	*entíto*	*encekût*	*ɔlmɛʊ́t*
Obl	*ɛŋkalém*	*ɛŋkárè*	*entitó*	*encékùt*	*ɔlmɛ́ʊt*

If we now look at the position of case markers with respect to the stem, we find that the well known 'suffixing preference' (see e.g. Hawkins and Cutler 1988) is in two-term case systems less prominent than in the languages of the world in general. According to Dryer (2005b), preposed case markers are found in less than 10 per cent of the languages with cases. However, among the languages with bicasual systems prefixal case markers are found in about 30 per cent of the linguistic groups, namely in Berber and Salish languages, Nias, and Movima. It is not at all obvious how these figures are to be interpreted, but they are nevertheless quite significant.

Turning to more complex issues, nominal paradigmatic structures observed in two-term case systems are often non-trivial. Certainly, the most common option is a separate exponent of (usually only Obl) case, invariable across different nominals; however, various deviations from this simple structure are attested. First of all, number and sometimes gender may be encoded cumulatively with case, as for example in the Indo-Iranian languages. Moreover, number may be expressed

separately, the case exponent being nevertheless sensitive to it, cf. the Khowar (Dardic, Èdel'man 1983: 212) paradigms in Table 47.5.

Table 47.5. Case and number exponents in Khowar

	'brother'		'son'	
	Sg	Pl	Sg	Pl
Dir	*brār*	*brār-gini*	*žau*	*žižau*
Obl	*brār-o*	*brār-gini-ān*	*žaw-o*	*žižaw-ān*

Various types of neutralization of categories are found, too. It is certainly common to have no case distinction in the plural, as for example in Yaqui and Aleut, but some languages (many Indo-Iranian throughout all nominals, as well as Old French in the subset of demonstratives) neutralize number in the Direct case, cf. the Kati (Nuristani, Èdel'man 1983: 60) paradigm in Table 47.6.

Table 47.6. Nominal paradigms in Kati

	'girl'		'man'	
	Sg	Pl	Sg	Pl
Dir	*ǰuk*		*manči*	
Obl	*ǰuka*	*ǰuko*	*manče*	*mančo*

In the Indo-European two-term case systems more 'exotic' patterns of syncretism are also found, such as the identity of Oblique Singular and Direct Plural, observed in many Indo-Iranian languages, cf. Table 47.7 with Pashto paradigms (Skjærvø 1989a: 390).

Table 47.7. Nominal paradigm in Pashto

	'Pashto'	
	Sg	Pl
Dir	*paṣtun*	*paṣtānə*
Obl	*paṣtānə*	*paṣtāno*

In Old French not only Oblique Singular and Direct Plural fall together, but quite often Direct Singular and Oblique Plural, too; this has led to a situation when four paradigmatic cells are covered with only one overt affix, cf. Table 47.8 (Pope 1934: 311).

Table 47.8. Nominal paradigm in Old French

	'wall'	
	Sg	Pl
Dir	*mur-s*	*mur*
Obl	*mur*	*mur-s*

To conclude this section, we may observe that the morphological make-up of bicasual systems has some peculiar characteristics which are seldom or never attested in richer case systems. Minimal systems are more prone to prefixal or non-concatenative case marking, as well as to unexpected patterns of paradigmatic neutralization.

47.5 Diachronic issues

There are two main questions concerning the diachrony of two-term case systems:

1. What are their possible diachronic sources?
2. What happens to them in the course of their history?

These issues will be briefly discussed in this section.

For the majority of the languages discussed in this paper, no written sources concerning their history are available. Therefore, any hypotheses about the origins of two-term case systems outside the Indo-European family are rather tentative. It is possible to discern the following two processes leading to a two-term case system: (i) reduction of a richer case system (as in the Indo-European languages); (ii) grammaticalization from other types of markers, such as adpositions (as in Salish, Movima, and probably Yimas) or demonstratives/articles (as in Berber, see Chaker 1988, and Ethiopian Semitic). The origins of tonal case systems of the Nilotic languages remain obscure.

The process of disintegration of polycasual systems into bicasual ones is well documented both for the Romance languages (e.g. Pope 1934) and for the Indo-Iranian languages (e.g. Kerimova and Rastorgueva 1975). It seems that various factors were interacting in the course of this development. Besides the obvious phonological erosion of case endings, syntactico-semantic processes were also of the utmost importance. The overlapping of the range of uses of the original cases led to their becoming interchangeable in many contexts, and to a decrease in the number of

grammatical oppositions. Numerous case syncretisms existing already in the older polycasual systems also facilitated the functional and morphological merger of originally different cases.

Of especial interest in this connection is the situation observed in Sogdian (Middle Eastern Iranian, Sims-Williams 1982), where an older system of five cases co-existed with a newer two-term case system. Such a rare situation became possible because of the so-called 'rhythmical law' (Tedesco 1926), which caused reduction of final syllables of the so-called 'heavy stems', with the 'light stems' remaining intact, cf. paradigms of masculine nouns of both types in the singular in Table 47.9 (Sims-Williams 1982: 67, 68).

Table 47.9. Nominal inflection in Sogdian (masculine, singular)

	'people' (light)		'day' (heavy)
Nom	*ram-i*	Dir	*mēθ*
Acc	*ram-u*	Obl	*mēθ-ī*
Gen	*ram-e*		
Loc	*ram-ya*		
Abl	*ram-a*		

However, the rhythmical law was not the only reason for the emergence of the bicasual subsystem in Sogdian; the two-term Dir ~ Obl distinction was observed throughout the feminine gender and in the plural of all types of nouns, where it appeared probably even prior to the operation of the rhythmical law. Thus the older polycasual system was a marked option restricted to a limited subset of nominals, in contrast to a newer two-term case system appearing elsewhere.

In the process of reduction of an older polycasual system into a bicasual one several older cases fall together thus forming a new case with a broader range of meanings. The resulting set of functions is not necessarily the simple unification of the uses of the predecessors of the new case, but the way the older system collapsed into a bicasual one crucially affects the structure of the latter. We have already seen (section 47.3) that the 'distributing' systems in the languages of the Hindukush-Pamir region reflect an earlier stage of development and retain the uses of older cases, while the majority of the 'dividing' systems in the Indo-Iranian languages must have redistributed the original case functions. If we compare the Indo-Iranian and the Old Romance two-term case systems, we see that the neutral alignment attested in the former and the predominant accusative alignment of the latter has clear diachronic origins. Indeed, during the disintegration of the Ancient Indo-Iranian polycasual system the original Nominative and Accusative fell together,

whereas in the course of the decline of the Latin case system the distinction between the Nominative and all other cases was usually retained (see e.g. Plank 1979).

Let us now turn to the fate of two-term case systems. Besides persisting for a long period without major changes (as probably is the case in the Nilotic languages), there are two major pathways of change:

1. A bicasual system may serve as a base for a newly grammaticalized polycasual system.
2. A bicasual system may disintegrate thus leaving the language without the case category altogether.

The first scenario is observed in the Indo-Aryan languages, where the primary function of the older inflectional two-term case system is to host a whole series of postpositions or even already bound case affixes (see Masica 1991: 238–48 for details). The postpositions usually attach to the Oblique case, and the new declensional system thus has two stems, one for the Direct (or Nominative) case, and another for all or most peripheral cases. The older Oblique itself in some languages may be used in isolation, thus becoming a new case, e.g. Accusative (as in Romani, Table 47.10, Wentzel 1980: 72), paradigmatically opposed to other cases formed upon its base, or, in others, functions only as the bound oblique stem, as in Gujarati (Table 47.11, Saveljeva 1965: 24–5).

The postpositional origin of such polycasual systems reveals itself in the behaviour of adjectives, which usually show up in a common Oblique form with all non-Nominative heads, cf. example (19) from Gujarati (Saveljeva 1965: 28).

(19) *glelɑ̃* *svapn-o-mɑ̃*
delirious:OBL.PL dream-OBL.PL-LOC
'in delirious dreams'

Finally, let us look at the situations when a two-term case system disintegrates. As in the fall of a rich case system, many factors play a role here. Sometimes it seems that Dir extends its usage, simply ousting Obl from its original contexts. Thus, in some Norwegian dialects Obl with nouns is used in the following functions: (i) as

Table 47.10. Nominal declension in Romani (North-Russian dialect)

	'pigeon'
Nom	*golúmbo*
Acc	*golumbó-s*
Loc	*golumbó-s-te*
Dat	*golumbó-s-ke*
Abl	*golumbó-s-tïr*
Ins	*golumbó-s-a*

Table 47.11. Nominal declension in Gujarati

	'dog'
Nom	*kūtro*
Gen	*kūtrā-n′*
Acc	*kūtrā-ne*
Ins	*kūtrā-e*
Abl	*kūtrā-thī*
Loc	*kūtrā-mã̄*

the Recipient of ditransitives; (ii) as a 'quirky' object of some non-canonically transitive bivalent predicates; (iii) as the object of some prepositions with a locational meaning ('live in the town'), contrasting with the allative meaning of Dir ('go to the town'); (iv) as a governed object of other prepositions. It appears (Sandøy 1996: 134) that Dir is taking over the more syntactic uses of Obl, viz. (ii) and (iv), while the latter persists in those contexts where it is semantically opposed to Dir, viz. (i) and (iii). Thus the range of functions Obl can encode gradually shrinks, and not in an unprincipled way.

A more complex situation was observed in late Medieval French (fourteenth–fifteenth centuries, see Laubscher 1921). Here the system was inherently unstable due to the typologically unusual markedness of the Direct case, coupled with its rather restricted range of functions, as well as the loss of phonologically overt markers. The less marked and more frequently used Obl began to gradually take over the uses of Dir, but there was a short period of free variation. As a result, in most cases Obl was retained and Dir lost, but sometimes they gave rise to different lexemes (e.g. *sire* < Dir in vocative function vs. *seigneur* < Obl), and in others it was Dir rather than Obl which persisted (e.g. *père* 'father').

47.6 Conclusions

In the preceding sections we have discussed various properties of two-term case systems, both functional and formal. Let us briefly review the principal points.

First of all, despite the seeming poverty of expressive means, two-term case systems more often than not cover quite a wide range of different semantic roles, not only the core grammatical relations, but also a more or less rich array of peripheral and circumstantial roles; 'poor' two-term case systems covering only the core relations are rather rare.

978–0–19–920647–6 48-Spencer-ch47 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 699 of 699 May 16, 2008 19:0

Second, markedness relations between cases in two-term case systems are usually driven not only by the typologically stable associations of marked expression with more 'peripheral' semantic roles and unmarked expression with more 'central' semantic roles, but also by the functional load of the cases. 'Marked nominative' systems, where the S/A relation is encoded by a marked case, while other semantic roles fall under the unmarked case, are quite widespread here.

Third, in two-term case systems such typologically rare case exponents as prefixes and proclitics or tonal modification are attested. Also, 'minimal' systems allow for paradigmatic structures which are usually not found in richer case systems, for example the Old French X-like neutralization of case and number.

Finally, from a diachronic perspective, two-term case systems constitute an important stage in the development of case in a language. They may be either the last stage of the reduction of an older polycasual system, as in the Romance and Iranian languages, or serve as the basis for a newer case system emerging via grammaticalization of adpositions usually attaching to the Oblique case as the stem.

CHAPTER 48

CASE IN IRANIAN

FROM REDUCTION AND LOSS TO INNOVATION AND RENEWAL*

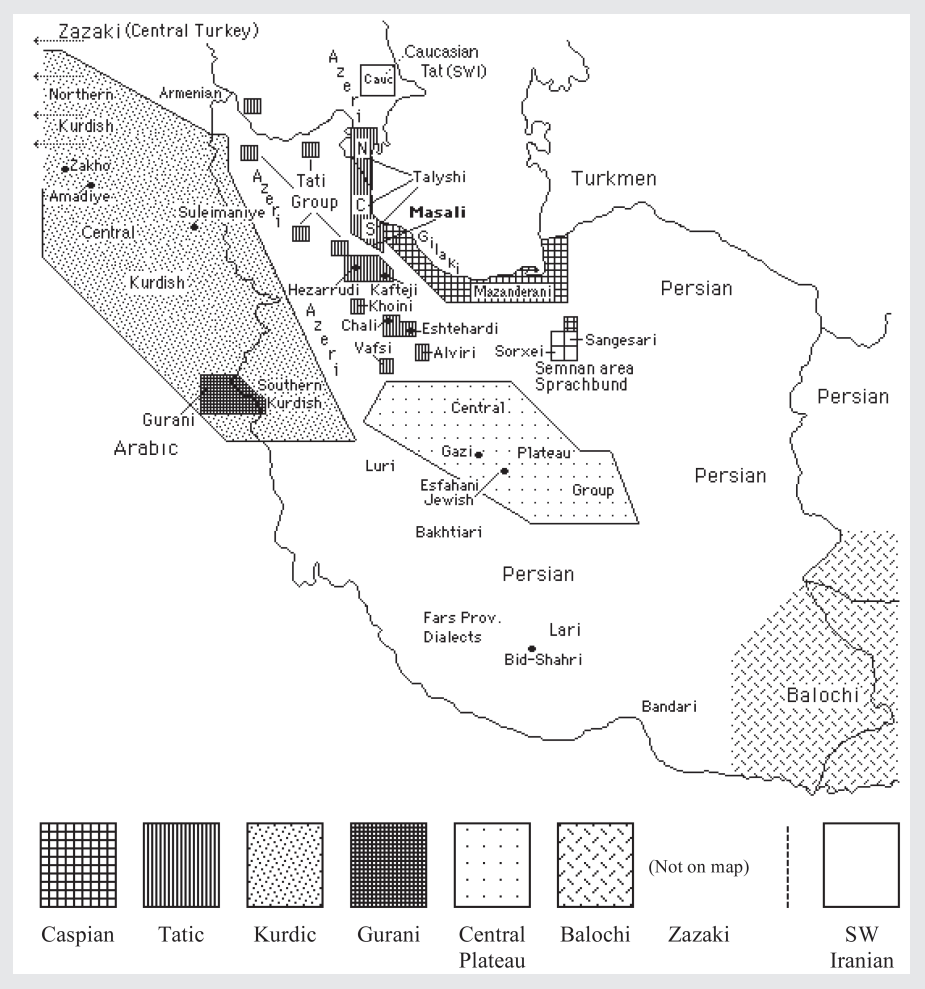
DONALD STILO

48.1 Introduction

Gathic Avestan, the earliest documented form of Old Iranian, had a rich case system typical of archaic Indo-European languages: three genders, eight cases in the singular, a six-term plural, and a four-term dual. Old Persian, a somewhat later stage of Old Iranian, had already begun some case syncretization with the old dative merging into the genitive. By the very earliest stages of Middle Persian and Parthian (the only attested West Iranian (WI) languages of the Middle period, see *CLI* (Schmitt 1989) for general information), gender and the dual were already completely lost and further stages of case syncretization yielded a two-case system: the null-marked Direct case vs. the Oblique case derived from the

* I wish to express my gratitude to the Swedish Collegium for Advanced Study, Uppsala, for its one-year invitation (2005–6) to join them as a fellow and for its generous support in the writing of the first few drafts of this article. I also would like to express my thanks to Andrej Malchukov, Geoff Haig, and the anonymous reviewer for their feedback and comments on this article.

Stilo's field notes were used for some or all of Vafsi, Gilaki, Talyshi, Kafteji, Razajerdi, Lari, and colloquial Persian. Some Balochi examples kindly provided by Prof. Carina Jahani, Uppsala University.



Map 48.1. Approximate locations of modern Western Iranian Languages

original genitive. By slightly later, better-known stages of these two languages, even this case distinction was ultimately lost. Various modern WI and East Iranian (EI) languages, however, still retain both gender and the inherited two-term case system, for example (WI) about thirty of the fifty or so Tati dialects (others have lost gender, as has the whole Talyshi group), Gorani, Zazaki/Dimli, Kurmanji/ Northern Kurdish, Semnani, etc.; (EI) Pashto. Some also preserve special forms for kinship terms. We can assume, therefore, that other unattested languages must have existed as predecessors to today's extant conservative languages. MacKenzie (1966: 3) states, for example, that modern Hawrāmī (Gorani) '… has the aspect of a somewhat archaic Middle Iranian dialect'. Thus some WI languages have maintained a stable two-term case system for at least 2,000 years. The Middle period of EI exhibited far

978–0–19–920647–6 49-Spencer-ch48 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 702 of 715 May 16, 2008 19:0

fewer syncretisms, but cases here were also eventually either reduced to two or completely eliminated by the modern period. Possible remnants of other cases found sporadically in modern Iranian may also be explained alternatively as innovations.

Reduction or loss of case throughout Iranian and the emergence of numerous innovations have led to a bewildering array of compensating strategies whose evolution can be characterized by the interplay of three axes/dimensions: the Reduction Axis, the Innovation Axis, and the Nominal-Pronominal Axis. This survey focuses primarily on WI.

48.2 The three axes

48.2.1 The Reduction Axis

The Reduction Axis reflects a progression from maximum *Retention* of the two-term system, e.g. Vafsi (Figure 48.1, Type 1) through varying degrees of further *Reduction* of the two-term type to complete *Loss* of all vestiges of case.

48.2.1.1 *Retention and the Scale of Reduction*

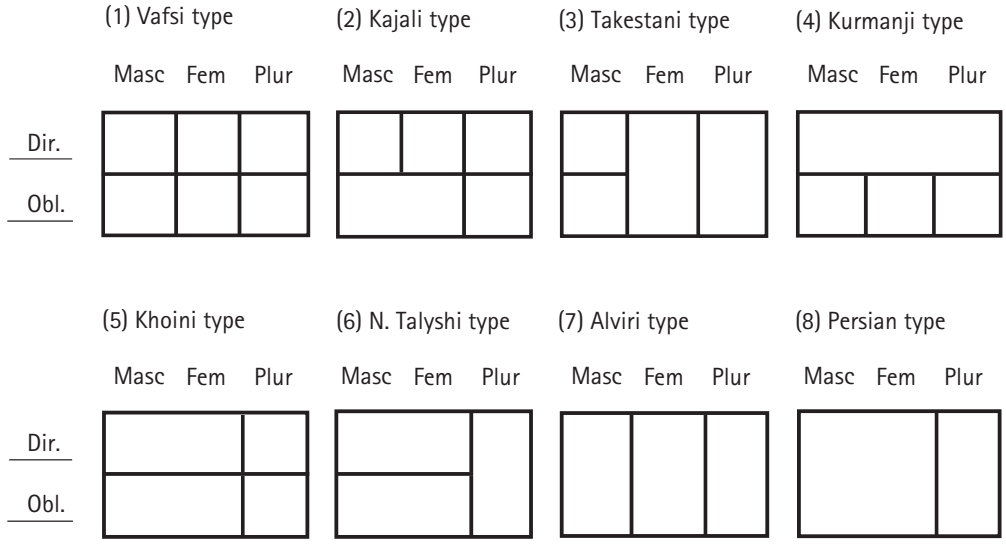
Retention of the most conservative two-term systems (separate direct–oblique cases for the masculine and feminine, kinship terms (see discussion below), and the plural) is found in languages such as Vafsi (Figure 48.1), Semnani, Kafteji (see (9) and (24)), Hezarrudi, Pashto, and others. With this level of conservatism as the point of departure, we find a whole range of gradual erosion of the case–gender–number systems in other modern Iranian languages. A sampling of this impoverishment is shown in the Scale of Reduction in Figure 48.1, followed by examples of six of these types.

In addition to the three major categories of masculine, feminine, and plural shown in the cline of case typologies, there are two additional inherited categories that would make the task of drawing up Figure 48.1 too complex to include:

- an Oblique II case marker, derived from Indo-European (and generalized in many languages, cf. Persian *bæradær*, 'brother', *xahær* 'sister', *madær* 'mother', *pedær* 'father', etc.), is used only for kinship terms (found in Tati, Zazaki, and in the Semnan area). The direct cases of these forms fall under Direct I masculine and feminine types but have a common Oblique II in -(*ǽ*)*r*. 'Brother' and 'sister' are:

	Kafteji		Vafsi		Esfarvarini (S. Tati)	
Dir.	bera-Ø	xóv-æ	bera-Ø	xák-æ	bera-Ø	xák-æ
Obl.	bera-r	xo-r	bera-r	xak-ár	bera-r	xak-ár

Typologies:



Some examples:[a]

	(1) Vafsi		
	Masc	Fem	Plur
Dir.	æsb-Ø	kǽrg-æ	ǽsb-e, kǽrg-e
Obl.	ǽsb-i	kærg-é	æsb-án, kærg-án

	(3) Takestani		
	Masc	Fem	Plur
	æsb-Ø	kǽrg-æ	æsb-ón, kærg-ón
	ǽsb-e		

	(5) Khoini		
	~~(Masc)~~	~~(Fem)~~	Plur
	æsb, kærg		æsb-é, kærg-é
	æsb-é, kærg-é		æsb-ón, kærg-ón

	(6) N. Talyshi		
	~~(Masc)~~	~~(Fem)~~	Plur
Dir.	asb-Ø, kag-Ø		asb-ón, kag-ón
Obl.	asb-í, kag-í		

	(7) Alviri		
	Masc	Fem	Plur[b]
	æsb	kǽrg-æ	æsb-há, kærg-há

	(8) Persian		
	~~(Masc)~~	~~(Fem)~~	Plur
	æsb, morq		æsb-an/æsb-há, morq-an/morq-ha

Figure 48.1. The Scale of Reduction to Loss of Case

[a] Examples given in Figure 48.1 mean 'horse' (*æsb*, m.), 'hen, chicken' (*kǽrgæ*, f.)

[b] While there are more than the eight types of case systems along the cline shown here, it is not always easy to find a perfect match for each category using only reflexes of original cases. Alviri, for example, does not fit Type 7 shown above exactly since the plural morpheme is not a reflex of one of the original plural case endings. The plural morpheme *-há* is in fact an innovation common to many WI languages, most notably modern Persian (see examples under (8) in this figure), most likely appearing in Alviri as a borrowing. This problem of exact fit with the typological categories was especially true for EI languages, except for Pashto which is a clear Type 1.

- Some languages, e.g. Vafsi and Gorani, have a feminine II type direct case ending in *-e* or *-i* which may be a reflex of an Old Iranian feminine type in *-ī̆*.

48.2.1.2 *Case loss*

Some WI have lost all case marking for the roles of agent, object, recipient, possessor, experiencer, location, goal, and temporal usages, innovating neither case-like enclitics nor adpositions to express them (but see Expansion below for other options), as in:

(1) *a esbá baqeban pür dē-m.* (Sorxei, Semnan area)
I dog gardener son gave-1SG
'I gave the dog to the gardener's son.' (Christensen 1935: 66)

Most languages, however, whether on the higher or the lower end of the Scale of Reduction in Fig. 48.1, have supplemented their systems with strategies of Innovation.

48.2.2 The Innovation Axis

The Innovation Axis includes processes of 1) *Expansion* (48.2.2.1): increased functions of the two inherited cases; extension of kinship oblique case forms to other animates or even inanimates, cf. Zazaki *jīp-er* 'jeep (F.OBL.)' (Paul 1998: 22); 2) *Renewal* (48.2.2.2): replenishing the system with case-like clitics; new case forms (e.g. locatives, feminine plural, etc.); 3) *Alternative strategies*: non-case-like strategies move in from outside the NP domain to shore up the diminished argument-marking system. (These alternative strategies are complex and interesting enough to be singled out for discussion in a separate section (48.3) just before the conclusion of the chapter.)

48.2.2.1 *Expansion*

The remaining two cases of the impoverished case system expand their scope to a surprisingly wide range of *polyfunctionality*, a common phenomenon in two-term systems (see Arkadjev, Chapter 47):

SUBJECT:	Dir	RECIPIENT:	Obl	GOAL:	Dir or Obl
AGENT:	Dir or Obl	POSSESSOR:	Obl	TEMPORAL:	Dir or Obl
OBJECT:	Dir or Obl	EXPERIENCER:	Obl	LOCATION:	Dir (or Obl)

Adpositions commonly mark GOAL, TEMP, and LOC, but are often optional, as in the Pāwä TEMP and GOAL (allative) uses, the Kurmanji LOC (but see also (49) below for caseless, adpositionless marking of locatives), and the Northern Talyshi TEMP:

978–0–19–920647–6 49-Spencer-ch48 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 705 of 715 May 16, 2008 19:0

(2) *šæw-**i** luā šæ'r-**i**-(æ).* (Pāwä, Gorani)
night-**OBL** went city-**OBL**-(DEF)
'In the evening he went to the city.' (Christensen 1921: 56)

(3) *am-ēt t-āxiv-īn āmēdī-ē.* (Amadiye dialect, Kurmanji)
we-TAM TAM-talk-1P PN-F.OBL
'We were talking in Amadiye (town).' (MacKenzie 1961: 198)

This wide expansion of case functions can lead to multiple uses of one case in the same sentence, as in this Vafsi example with five obliques:

(4) *æhmǽd-**i** ær-go vaár-**i** mæhmúd-**i** ǽsb-**i***
PN-**M.OBL** TAM-want spring-**M.OBL** PN-**M.OBL** horse-**M.OBL**
EXPERIENCER TEMPORAL POSSESSOR OBJECT
*há-do-æ jævád-**i**.*
PREV-give-3SG PN-**M.OBL**
RECIPIENT
'In spring Ahmad wants to give Mahmud's horse to Javad.'
(Stilo, field notes)

Two prevalent Iranian marking splits, ergativity and Differential Object Marking (see *Renewal*), further diversify the uses of these two cases.

Ergativity and the Double Oblique

Using this two-term case system, many ergative Iranian languages show a tense-based split of Nominative–Accusative alignment (direct and oblique cases, respectively) in the present system vs. Ergative–Absolutive alignment (oblique and direct cases, respectively) in the past system ('past' generally defined morphologically as all tenses based on the past stem of transitive verbs):

(5) a. *sābir-Ø ē hawāl-ā uškin-īt.* (Kharan, Balochi)
PN-DIR this news-OBL hear-3SG
'Sabir hears this news.' (Korn, forthcoming, citing Elfenbein 1990/I: 62)
b. *sābir-ā ē hawāl-Ø uškit.*
PN-OBL this news-DIR heard
'Sabir heard this news.' (Carina Jahani, p.c.)

(6) a. *az tolïši zïvon-í o-bæ-müt-é=m* (N. Talyshi, Tatic)
I.DIR Talyshi language-OBL PREV-TAM-learn-TAM=1S
'I'll learn Talyshi.'
b. *mï tolïši zïvon-Ø=ïm o-mű́t-ï*
I.OBL Talyshi language-DIR-1SG PREV-learn-TAM
'I learned Talyshi.' (Stilo, field notes)

(7) a. *vī ṯayra-y nā-kiṙ-ī* (Akre dialect, Kurmanji)
this.M.OBL bird-M.OBL NEG.TAM-buy-2S
'Won't you buy this bird?' (MacKenzie 1961: 155)

b. *aḥmad-ī ḥākim t=sindōq-ē nā*
PN-M.OBL governor in=chest-F.OBL placed
'Ahmad put the governor into the chest.' (MacKenzie 1961: 194)

In addition to the type of canonical ergative alignment in the past tenses of conservative Iranian languages where we find the agent in the oblique, the object in the direct, and the verb agreeing with the object, there is a wide range of non-canonical ergative-derived constructions in various Iranian languages that deviate from this norm (see also Table 48.1). Some WI and EI languages, for example, have 'Double Obliques' with both agent and object in the oblique in past tenses:

(8) *tæmen æsb-i=m bǽ-di.* (Vafsi, Tati)
I.OBL horse-M.OBL=1SG TAM-saw
'I saw the horse.' (Stilo, field notes)

(9) *æmǽd-e čem bera-r vend* (Kafteji, Tati)
PN-M.OBL I.POSS brother-KIN.OBL saw
'Ahmad saw my brother.' (Stilo, field notes)

(10) *bačakk-ā watī dantān-ā̃* prōšt (W. Balochi)
boy-OBL own teeth-PL.OBL break.PAST
'The boy broke his teeth.' (Korn, forthcoming)

(11) *mu dum kitōb=um χēyt.* (Roshani, EI)
I-OBL [this.OBL book=1SG] read.PAST
'I read this book.' (Payne 1989: 439)

Additional possibilities for core argument marking with the direct and oblique cases are given below under Differential Object Marking for which we must first introduce accusative marking under Renewal strategies.

48.2.2.2 *Renewal strategies*

Renewal strategies introduce innovations that enrich both caseless as well as two-term systems. The discussion below covers: (i) case-like accusative markers; (ii) the *ezafe* construction; (iii) singular cases suffixed to plurals in a somewhat agglutinative fashion; (iv) renewal by areal analogy; and (v) case-like adpositional clitics.

(i) Renewal strategy one: *=ra*. The Persian enclitic *=ra* (colloquial = (*r*)*o*), derived from the Old Persian postposition *rādiy* 'for (the sake of)' – now performing only a syntactic function devoid of semantic content – is used to mark specific direct objects. In older forms of New Persian, *=ra* marked various core and non-core arguments as it still does, for example, for the Caucasian Tat (SWI, unrelated to Tatic) direct object (12), indirect object (13), experiencer (14), and possessor of 'to have' (15), but also including general possession (16):

(12) *dowšán=æ gűft=im.*
rabbit=RA caught=1PL
'We caught the rabbit.' (Grjunberg 1963: 23)

(13) *in ædæmi-ha=ræ mu-kutrǽn=üm.*
this person-PL=RA TAM-say=1SG
'I tell these people.' (Grjunberg 1963: 123)

(14) *xünűk=ü tű=ræ?*
cold=is you=RA
'Are you cold?' (Grjunberg 1963: 107)

(15) *ispiškæ hist tű=ræ? nist mǽn=æ, æhmǽd=æ hist.*
match is you=RA is.NEG I=RA PN=RA is
'Do you have any matches? I don't, Ahmad does.' (Grjunberg 1963: 25)

(16) *xunǽ=ræ duvár=i*
house=RA wall=3SG.POSS
'the wall of the house' (Grjunberg 1963: 26)

In Caspian languages, =(*r*)*a* marks only some of the above range of syntactic roles – direct objects, indirect objects, and experiencers, as in:

(17) *mən dastán-a həsǿn-a bu-gúft-əm. án=a xoš*
I story=RA PN=RA TAM-said-1SG he=RA pleasant
OBJECT RECIPIENT EXPERIENCER
b-amo. (Gilaki, Caspian)
TAM-came
'I told Hasan the story. He liked it.' (Stilo 2001: 662, and field notes)

DOM: Differential Object Marking. While early New Persian =*ra* covered most of the above usages, Modern Persian uses =*ra* exclusively to mark specific direct objects (18). Object marking in Caucasian Tat, Caspian languages, many SWI, and others is also linked to specificity and referentiality. (For generalities, see also Malchukov and de Swart, Chapter 22, as well as Bossong 1985d for DOM in Iranian languages.)

(18) *nun=o kære xæríd-æm. kærǽ=ro koja*
bread=and butter bought-1SG butter=RA where
gozašt-æm? (Colloquial Persian)
put.PAST-1SG
'I bought bread and *butter*. Where did I put *the butter*?'

Many Central Plateau Dialects (henceforth CPD) have borrowed Modern Persian =*ra* and a copy of its usage ((19), (48b)), while some Iranian languages have developed their own corresponding *ra*-like accusative marker from other adpositions such as a locative in Sangesari =*d*(*e*) (20):

978–0–19–920647–6 49-Spencer-ch48 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 708 of 715 May 16, 2008 19:0

(19) *gorg-á guθpand-á=ra b-ešun-xo.* (Esfahani Jewish, CPD)
wolf-PL sheep-PL=RA TAM-3P-ate
'The wolves ate the sheep (pl.).' (Stilo, in press a)

(20) *atæš=de bə-kóšt.* (Sangesari, Semnan area)
fire=ACC TAM-killed
'He extinguished the fire.' (Christensen 1935: 136)

These accusative markers of adpositional origin often retain their spatial-locational senses, as the Sangesari *=de* 'at'. Munji *va* and Persian *=ra*, however, only have a syntactic function as accusative marking, lacking any additional intrinsic semantic content. Note that *=ra* in most NWI never evolved into an accusative marker and still retains a benefactive sense, but in many languages of northern Iran, this postposition is often polysemous (Vafsi: benefactive, instrumental/comitative; Razajerdi (Tati): benefactive, ablative; Chali: benefactive, allative, see example 43).

The restriction of accusative *=ra* to specific objects was not always true of Persian. Paul (2002) points out that *=ra* in early Judaeo-Persian texts (approximately eleventh century AD) was triggered not by specificity (21), but by animacy (22) (see also Vafsi example (29) below, as well as Korn (forthcoming) for a discussion of the role of animacy in object marking in Balochi, and Middle Iranian Parthian and Bactrian):

(21) *cwn 'yn skwn byšnyd.* (Paul (2002) transliteration)
čōn ēn saxwan bē-šnīd. (approximate transcription)
when this speech TAM-heard
'When he heard this speech'

(22) *kwd' mn r' prstyd.*
xudā man rā firistīd.
God I RA sent
'God sent me.'

Languages that use the oblique case for accusative marking also exhibit DOM based in specificity:

(23) a. *ši-n-ē resen-Ø ā-n-ē.* (Siverek dialect, Zazaki)
go-TAM-3PL rope-M.DIR bring-TAM-3PL
'They go bring a rope.'
b. *resen-ī nē-ver-dā-n-ē jēr.*
rope-M.OBL NEG-PREV-give-TAM-3PL down
'They don't lower the rope down.' (Paul 1998: 16)

(24) a. *gel-Ø á-xšuj!* (Kafteji, Tati)
clay-M.DIR PREV-knead
'Knead (some) clay!'

978–0–19–920647–6 49-Spencer-ch48 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 709 of 715 May 16, 2008 19:0

b. *árd-e* *á-xšuj!*
flour-M.OBL PREVERB-knead
'Knead the flour!' (Stilo, field notes)

An innovated case formed with *=(r)a* has been generalized as a multifaceted oblique in Balochi (see also (5)), but DOM still applies here:

(25) a. *sēb-Ø* *b-ōr!* (Sarbaz dialect, Iranian Balochi)
apple-DIR TAM-eat
'Eat apples!'
b. *ē* *sēb-ā* *b-ōr!*
this apple-OBL TAM-eat
'Eat this apple!' (Carina Jahani, p.c.)

Since all forms of Kurmanji lack DOM, obliques mark even non-specific objects:

(26) *čand* *ṙôž=a* *az=ē* *nān-ī* *da-m=a*
how.many day=BE.3SG I.DIR=TAM bread-M.OBL give-1SG=ALL
hinga? (Amadiye, Kurmanji)
you.PL.OBL
'How many days is it that I am giving you bread?' (MacKenzie 1961: 205)

(27) *kāγaz-ā* *t-nivīs-īt.* (Akre dialect, Kurmanji)
letter-PL.OBL TAM-write-3SG
'He is writing letters/he writes letters.' (MacKenzie 1961: 155)

Many EI languages mark both definite and indefinite objects and thus also lack DOM:

(28) *ta* *va* *kə́rya* *āvə́rγ-ət.* (Munji, EI)
you ACC.MARKER hen bought-2SG
'You bought a hen.' (Skjærvø 1989b: 415–16)

Vafsi restricts DOM to *specific*, *animate* objects in the present (see (4)) and in the past Double Oblique construction (see also (8) and Table 1.1). Since non-specific or inanimate objects are not oblique-marked, the structure is no longer double oblique:

(29) a. *tæmen* *æsbǽ-Ø=m* *bǽ-diæ.*
I-OBL dog-M.DIR=1S TAM-saw
'I saw a dog.'
b. *tæmen* *æsbǽ-y=m* *bǽ-diæ.*
I-OBL dog-M.OBL=1SG TAM-saw
'I saw **the dog**.' (Stilo, field notes)

The mirror-image may appear in the present: both agent and non-specific and/or inanimate object are in the direct case.

Table 48.1. Various configurations of Direct/Oblique marking of A/O in Vafsi

Case of A/O (+ Word Order of A,O)	Case of A/O	+/−	Object Marked (+) or Unmarked (−) (type of Object)	Pres/Past Tenses	Name of Construction (Past Tense Constructions in Descending Frequency)
$Direct_A$	$Oblique_O$	+	Specific, Animate only	Present	Nominative–Accusative
$Direct_A$	$Direct_O$	−	Non-Specific and/or Inanimate	Present	(DOM version of above)
$Oblique_A$	$Oblique_O$	+	Specific, Animate only	Past	Ergative–Accusative (or 'Double Oblique')
$Oblique_A$	$Direct_O$	−	Non-Specific and/or Inanimate	Past	(DOM version of above)
$Direct_O$	$Oblique_A$	+	Specific, Animate only?	Past	Absolutive–Ergative
$Oblique_O$ (rare construction)	$Oblique_A$	+	Specific, Animate only (probably only +)	Past	Accusative–Ergative (or 'Double Oblique II')

Vafsi, however, also has two additional, much less commonly used past tense constructions: an ergative $Object_{DIRECT}$–$Agent_{OBLIQUE}$ and an $Object_{OBLIQUE}$–$Agent_{OBLIQUE}$, both of which are obligatorily OSV. When we put these various constructions together and add the possibilities provided by DOM, we find a surprising array of choices for marking core arguments with just two cases in this language (Table 48.1).

While a general survey of case in Iranian is not the appropriate place to present a full discussion of Ergativity constructions in one language, it is relevant to show the intricate range of case usage that has evolved here. A fuller description of the above constructions and how they are distinguished from one another also involves, aside from tense, the various options for how and where the agreement with agent and/or object is encoded in the past tenses in this language (see section 48.3.1 below).

Some languages, such as Sorxei (see (1)), some CPD, and many SWI languages simpy have no DOM or any accusative marking whatsoever:

(30) *čína mehmon-á-tu to-æ-šúd* (Bid-Shahri, Lari, SWI)
how guest-PL-2PL 2PL-TAM-seated
'How did you (pl.) seat your guests?' (Stilo, field notes)

(ii) **Renewal strategy two**: A major strategy for case-like renewal is the *ezafe* construction, which consists of an unstressed suffix (the '*ezafe*') on a head noun connecting it to adjuncts, primarily possessors and adjectives (exclusively head–adjunct order), with three subtypes evolving in different languages:

1) an exclusively head-marking *ezafe* with one single variant (Persian, most SWI, some Central Kurdish, some CPD):

(31) *pedǽr-e æhmæd* (Persian)
father-EZAFE PN
'Ahmad's father'

978–0–19–920647–6 49-Spencer-ch48 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 711 of 715 May 16, 2008 19:0

(32) *bāwk-ī* *kuṙ* (Suleimaniye dialect, C. Kurdish)
father-EZAFE boy
'the boy's father'

2) a double-marking type with both an *ezafe* on the head and an oblique case on the possessor adjunct, further subdividing into:

a) one single *ezafe* variant (Gorani, Vafsi):

(33) *gōs-u* *dēw-ā* (Pāwä dialect, Gorani)
ear-EZAFE demon-PL.OBL
'the demons' ear(s)' (Christensen 1921: 57)

b) multiple *ezafe* variants before nominal adjuncts reflecting the number, gender, and the syntactic function of the head noun, for example three forms in Kurmanji and some nine forms in Zazaki.

On the other hand, some languages have eliminated all genitival marking, opting for simple juxtaposition, for example Sorxei (Genitive-head, see (1)) and many CPD (head–Genitive):

(34) *bær* *kie* *hæsir-baf* (Gazi, CPD)
door house mat-weaver
'the door of the mat weaver's house' (Stilo, in press b)

(35) *bele-tær-á* *mæhælle* (Esfahani Jewish, CPD)
big-COMPARATIVE-PL neighborhood
'the elders of the neighbourhood.' (Stilo, in press a)

(iii) **Renewal strategy three**: agglutinative analogy. While some languages lost the direct/oblique distinction in the plural, Caspian languages, Sangesari, and some EI have renewed the distinction by adding the singular oblique morpheme to the plural (originally the oblique plural, now generalized) in a somewhat agglutinative fashion:

(36) Oblique Sg. Oblique Pl. (Rashti, Gilaki)
kor-ə *mar* *kor-an-ə* *mar*
girl-OBL mother girl-PL-OBL mother
'the girl's mother' 'the girls' mother' (Stilo, field notes)

A similar agglutinative pattern exists in the Balochi plural where the inherited oblique (*-ay* ~ *-ī*) is restricted to genitival function (see (25) for the modern Balochi oblique):

(37) *mn-ī* *guhār-ay* *nām* *mn-ī* *guhār-ān-ī* *nām* (Sarbaz Balochi)
I-GEN sister-GEN name I-GEN sister-PL-GEN name
'my sister's name' 'my sisters' name' (Carina Jahani, p.c.)

(iv) **Renewal strategy four**: code-copying via language contact. The richest innovated case system within Iranian is seen in Ossetic (EI, spoken in Georgia and

978–0–19–920647–6 49-Spencer-ch48 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 712 of 715 May 16, 2008 19:0

the North Caucasus area of Russia), which has developed a nine-term, somewhat agglutinative-type paradigm, similar to neighbouring Georgian, Armenian, and Azerbaijani:

(38)

Singular ('head')		Plural ('heads')	(Iron, Ossetic)
Nominative	*sær-Ø*	*sær-tæ-Ø*	
Genitive	*sær-ï*	*sær-t-ï*	
Dative	*sær-æn*	*sær-t-æn*	
Allative	*sær-mæ*	*sær-t-æm*	
Ablative	*sær-æy*	*sær-t-æy*	
Locative	*sær-ï*	*sær-t-ï*	
Superessive	*sær-ïl*	*sær-t-ïl*	
Equative	*sær-aw*	*sær-t-aw*	
Comitative	*sær-imæ*	*sær-t-imæ*	(Abaev 1964: 19–20)

(v) **Renewal strategy five**: adpositions and case-like clitics. All Iranian languages have simply dealt with the loss or reduction of cases with the well-known process of replacing cases with adpositions. Throughout the family we find postpositions (including enclitics), prepositions (including proclitics) and circumpositions (see Stilo 2006). Some of these adpositional clitics, however, while not exactly true cases, act as a type of renewal of case in the form of 'case-like' enclitics.

(39) *ti per=jə* (Rashti)
your father=from
'from your father!' (Stilo 2001: 662)

(40) *zærf=dæ* (Eshtehardi, Tati)
dish=LOC
'in the dish' (Yarshater 1969: 126)

(41) *hæsæn=endu* (Chali, Tati)
PN=with
'with Hassan' (Yarshater 1969: 123)

Many enclitic adpositions, however, are not case-like since they themselves, in some languages, require an oblique case on the accompanying nouns:

(42) *bān-ān=di* (Zazaki)
house-PL.OBL=LOC
'in houses' (Paul 1998: 25)

(43) *hæsæn-e=ra* (Chali)
PN-M.OBL=RA
'for Hassan' (Yarshater 1969: 123)

48.2.3 The noun–pronoun axis

Some languages retain archaic forms or innovate new cases exclusively for the pronominal subsystem or, occasionally, only for some persons (e.g. Sorxei 1st and 3rd sg.). Pronominal systems are described only briefly here (but see Haig, forthcoming).

Hezarrudi (Tati) presents a fascinating example of pronominal case enrichment: an original tense-based split ergative opposition has expanded to a four-way opposition for S/A/O in some persons, for example 1SG direct *æz* (subject, present agent), (innovated) accusative *ædem* (present object), oblique *men* (past agent), (innovated) objective *æčem* (past object), e.g.:

(44) *eštæ* *bera* *æčem=eš* *be-zze*
you.GEN brother 1SG.OBJECTIVE=3SG TAM-hit.PAST
'Your brother hit me.' (Yarshater 1970: 460)

The innovated formants *æd-* and *æč-* derive from the Old Iranian prepositions *hada* 'with' and *hača* 'from,' respectively.

Many languages have innovated non-clitic pronominal genitives, for example Gilaki *mi* 'my', *ti* 'your', etc. or Tatic and some EI languages (e.g. Pashto) also with a formant derived from *hača* 'from' prefixed to the oblique pronoun, e.g. Hezarrudi *če-men* 'my' (a fifth pronominal case for this language). In Semnani, some Tati languages, and others, this innovated possessive then supplanted the original oblique case for marking A/O, expanding the new pronominal possessive into a general oblique, thus reverting to a two-term system – a type of Decay > Renewal > Decay.

Northern Talyshi, which already had a direct–oblique case distinction in a typically Iranian tense-based alignment split for nouns, has also innovated an accusative case only for pronouns irrespective of tense, e.g. direct/oblique/genitive/accusative (1st sg.) *az/mï(n)/čïmïn/mïní*, (2nd sg.) *tï/tï/ištï/tïní*, etc.

(45) *æy* *mïní* *voγánde* (Northern Talyshi)
he.OBL I.ACC send.PAST
'He sent me.' (Miller 1953: 119)

48.3 Alternative strategies

In addition to the above strategies, there are various non-case-like innovations helping to disambiguate A and O in languages with case-impoverished systems.

978–0–19–920647–6 49-Spencer-ch48 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 714 of 715 May 16, 2008 19:0

48.3.1 Floating clitics

In many Iranian languages, Person Agreement Markers (PAMs) encoding past agents in the verb are mobile clitics that commonly 'float' leftwards attaching to various other hosts ((6b), (8), (11), (29), (44)). Gazi, however, has developed a typologically unusual situation: although the clitic encoding agent may optionally remain on the verb or float leftwards, it quite consistently attaches to the object of the verb, and thus performs the double function of encoding agent while simultaneously, *by position,* identifying its host as the object. When the object moves leftwards, triggered by information structure, the agent clitic moves with it, but the clitic does not obligatorily move to second position. That is, despite the few examples below that seem to be to the contrary, there is no Wackernagel rule in Gazi:

(46) a. *bé=ž=irint* (Gazi, CPD)
TAM-3SG-bought
'He bought'
b. *kie=ž b-írint.*
house=3SG TAM-bought
'He bought *a/the house*.'
c. *ye kíe bæle=ž æz mo b-írint-e=bo.*
one house big=3SG from I TAM-bought-PRF=AUX.PAST
'He had bought *a big house* from me.' (Stilo, in press b)
d. *ye šal yæzzi nu=ž taže bičare b-írint-e=bo.*
[one scarf Yazdi new]=3SG just poor TAM-bought-PRF=AUX.PAST
[OBJECT]=AGENT.PAM ADV AGENT VERB
'The poor (guy) had just bought a new Yazdi scarf.' (Stilo, in press b)

This grammaticalized double function of the clitic in Gazi is clearly a *tendency* in some languages, e.g. Vafsi, Esfahani Jewish, possibly Central Kurdish and Iranian Balochi (Korn, forthcoming). Iranian Balochi needs more investigation to see to what extent this pattern has become consistently grammaticalized:

(47) *alī-Ø hasan-Ø=ī zat.* (Lashari, Iranian Balochi)
PN-DIR PN-DIR=3SG struck
'Ali hit Hasan.' (Korn, forthcoming)

In some CPD, objects of past verbs are optionally marked either with *=ra* (+specific) or with the agent PAMs fronted from the verb (±specific), but are never doubly marked and thus have variable DOM. That is, while the *=RA* option only marks specific direct objects, the fronting of the agent clitic to the object as host is indifferent to DOM.

(48) a. *ow-š bə-xo*
water=3SG TAM-drank
'He drank (some/the) water.'

b. *ow=ra* *bé-š-xo* (Hamadani Jewish, CPD)
water=RA TAM-3SG-drank
'He drank the water.' (Stilo 2003: 624)

48.3.2 Word order

Word order constraints often contribute to disambiguating arguments that share the same case (or caseless) forms. In some languages, certain arguments are usually distinguished by word order and adpositions are optional. See for example the recipient in (4) and GOAL vs. LOC in:

(49) a. *færda* *míræm* *mædrese.* (Colloquial Persian)
tomorrow go.1SG school
'I'm going to school tomorrow.'
b. *færda* *mædrese* *mímunæm.*
tomorrow school stay.1SG
'I'll stay at school tomorrow.'

This example is not meant to imply that this word order contrast in colloquial Persian, as well as most languages of western Iran, has become fully grammaticalized. It is only a common pattern in the language, and neither these word order patterns nor the deletion of the adpositions should be considered rigidly fixed.

48.4 Conclusion

Iranian languages have generally supplemented the impoverishment of case systems with strategies of Expansion-Renewal-Innovation. For example, Vafsi (first, most conservative type in Figure 48.1) has Retention + Renewal (DOM, *ezafe*) + Expansion (wide use of the oblique) + Innovation (four constructions for marking core arguments, PAM clitic mobility). Kurmanji has an intermediate level of Retention (gender or plural not expressed overtly in direct, only in oblique) + Renewal (extensive range of *ezafes*) + Expansion (wide use of the oblique). At the opposite end of the spectrum (types (7) and (8) in Figure 48.1), Sorxei and some CPD lack case, accusative marking, the *ezafe*, or other Genitival marking and rely primarily on juxtaposition and word order. Typologically expected developments for marking all non-core arguments with adpositions, however, are also common as replacements or supplements throughout Iranian.

CHAPTER 49

FROM SYNTHETIC TO ANALYTIC CASE

VARIATION IN SOUTH SLAVIC DIALECTS

ANDREJ N. SOBOLEV

49.1 Introduction

There are two different classifications of South Slavic languages (Slovenian, Croatian, Bosnian, Montenegrin, Serbian, Macedonian, and Bulgarian) available which consider them within a single dialect continuum stretching from the Slovenian Alps to the Black Sea, viz. one genetic and the other typological (Birnbaum 1965). The first classification opposes two groups, viz. historically Western and Eastern–Southern Slavic dialects (Sobolev 1998). Roughly, the former is represented by Slovene, Croatian, and Serbian and the latter by Bulgarian and Macedonian. The typological classification divides these languages into two groups as well. The first comprises the dialects that retain the historical synthetic noun inflection typical of Slavic generally, as reconstructed for proto-Slavic or attested in Old Church Slavonic, Russian, or Polish. The dialects that gave up synthetic morphology in favour of analytic constructions belong to the second group (Hinrichs 2000). The nominal paradigm of a synthetic Southern Slavic language may include up to six synthetic cases: nominative, accusative, dative, genitive, instrumental, and locative;

vocative case is not considered here. Synthetic systems are therefore archaic; a typical representative for this group is a system of Shtokavian dialects (and/or the Serbian literary standard in general, Stevanović 1975), e.g. *knjig-a momk-a* (book-NOM boy-GEN) '(the) boy's book'. The innovative group comprises the languages with analytical systems, such as Bulgarian, Macedonian, and Eastern Serbian dialects. These languages have lost case inflection. Normally there will be either one nominal form labelled *general case*, or two, the *general direct case* (DIR, originally a nominative), and *general oblique case* (GENERALOBL, originally an accusative). Syntactic roles which historically were realized by synthetic cases are normally marked in these systems by word-order, intonation, prepositions, and in certain Balkan languages by shortened forms of personal pronouns (to denote a (in)direct object or, rarely, possession) as well. The typical representatives of this group are eastern Bulgarian dialects and/or the Bulgarian literary standard (Gramatika 1982–1983), e.g. *knig-a-ta na momč-e-to* (book-DIR-DEF on(PREP) boy-DIR-DEF)' '(the) boy's book'. Serbian Kosovo–Resava dialects belong to the Western-Northern periphery of the analytical languages, since they occupy a transitional stage from synthetic to analytic organization (Cyhun 1981; Ivić 2001; Miloradović 2003).

49.2 Balkanization and analyticization

The change from synthetic to analytic patterns in certain South Slavic dialects is the result of many centuries of contact with the neighbouring Balkan languages (particularly East Romanian, Albanian, and Greek). This development began in the sixth–seventh centuries AD and is known as the 'Balkanization' of these languages (as they became the South Slav member of the Balkan Sprachbund).[1] The balkanized South Slavic languages are also known as Balkan Slavic languages. The process of balkanization may be illustrated by the development of the Slavic instrumental case. The set of meanings of a non-predicative verbal prepositionless instrumental case in South Slavic may be deduced from the following samples of the earliest (tenth-century) written language – Old Church Slavonic (Eastern South-Slavonic, with Macedono-Bulgarian features at its core). Notwithstanding alternations in the selection of prepositional constructions (e.g. *sъ* + instrumental in comitative meaning, *otъ* + genitive case in agentive meaning), it seems probable that at the earliest stages of Southern Slavic, the instrumental case was an indivisible category with a defined set of meanings (1a – instrumental with the meaning 'instrument'

[1] On the Balkan Sprachbund and so-called Balkan linguistic features see e.g. Sandfeld (1930), Koneski et al. (1968), Mladenov and Steinke (1978); Popov (1984); Hinrichs (1999), Mišeska-Tomić (2004).

978–0–19–920647–6 50-Spencer-ch49 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 718 of 729 May 16, 2008 19:0

with transitive verbs; 1b – the same but with intransitive verbs,[2] 1c – instrumental of means; 1d – comitative as modifier; 1e – agent in a passive construction; 1f – instrumental of cause; 1g – instrumental of aggregate; 1h – instrumental of limitation with intransitive verbs; 1i – instrumental of place; 1j – instrumental of time):

(1) a. *pilojǫ prětiraaáe*
saw.INS he.cut
'(he) cut with a saw'
b. *i rǫkama plešte likujǫ*
and hands.INS applauding I.triumph
'and clapping my hands, I triumph'
c. *idošę korablemь*
they.went ship.INS
'they travelled by ship'
d. *rače drьzostьjǫ*[3]
he.said imprudence.INS
'(he) said imprudently'
e. *ljubimъ jestъ bogomъ*
loved.PASS is God.INS
'is loved by God'
f. *azъ že sъde gladomь gьibnǫ*[4]
I EMPH starvation.INS I.perish
'as for me, I am perishing here of starvation'
g. *krovi strujami tekušti*
blood stream.INS.PL flowing
'blood flowing in streams'
h. *onъ tělomъ umrě*
he body.INS died
'his body died' (lit. 'he died with his body')
i. *pride k nimъ polemъ*
came to them field.INS
'came to them across the field'
j. *učenici ego noštijǫ prišьdъše*
disciples his night.INS having.come
'his disciples, having come at night'

[2] That is, the use of the instrumental case in the object function, cf. Russian *brosit' kamen'* 'throw.INF stone.ACC' ~ *brosit' kamnem* 'throw.INF stone.INS' 'to throw a stone', Russian *maxat' rukami* 'wave.INF arm.INS.PL' ~ Bulgarian dialect *m'axʌ rʌk'ite* 'wave.INF arm.GENERALOBL.PL.DEF' 'to wave one's arms'.

[3] Cf. Greek ἔφη μετὰ παρρησίας. For phonological reasons examples of the prepositionless instrumental in the sociative meaning in Old Church Slavonic are not reliable.

[4] Cf. Old Russian *razbolěsja volodimerъ očima* 'became.ill Volodimer eyes.INS' '(lit.) Volodimer became ill in the eyes'.

In general, this category in Slavic tends to expand (possibly by analogy) from the functional/semantic centre of the 'instrumentality' field to the periphery, applying to a greater variety of verb types. The most visible feature of this development in the Balkan Slavic languages is the neutralization of the instrumental and comitative (normally in the constructions *s* 'with' + instrumental or *s* 'with' + general oblique case[5]) following the Balkan model (which is also known from other European languages, Sobolev 2005: nos. 75, 77). Another peculiarity of the Balkan and Balkan Slavic languages can be seen in the generalization of the historical ablative construction with the preposition *ot* 'from; of' into the marker of the passive agent (Sobolev 2005: no. 67). This feature sharply distinguishes the Balkan linguistic type from that of Northern Slavic, which contrasts instrumental and comitative (cf. Russian *režu nožom* 'I.cut knife.INS' 'I cut with a knife' ∼ *idu s Ivanom* 'I.go with Ivan.INS' 'I go with Ivan', cf. 1a and 1d), and in which the instrumental is used to mark the agent in a passive construction (cf. Russian *dom postroen rabočimi* 'house.NOM built.PASS.PTCP workers.INS' 'The house has been built by workers', cf. 1e). The most important Balkan trait in the development of the case system in Southern Slavic languages, however, is the change from synthetic to analytic, that is the change from marking syntactic functions with inflections to marking them by prepositions. This process is reflected to varying degrees in East Serbian dialects (Sobolev 1991a; Miloradović and Greenberg 2001; inter alia): *režem nožem* 'I.cut knife.INS' > *režem s nožem* 'I.cut with knife.INS' > *režem s nož* 'I.cut with knife.GENERALOBL' 'I cut with a knife'. Changes in encoding type and neutralization of the instrumental/comitative opposition does not necessarily lead to a wholesale redistribution of this function, and its set of meanings (apart from agentive, local, and temporal meanings usually) may even coincide with that of the original Slavic instrumental.

Functions/meanings corresponding to peripheral domains of the Slavic instrumental (such as agentive meaning, cause, and limit) are very often represented by constructions with an ablative–agentive preposition *ot*, as in Macedonian dialects:

(2) a. *beše 'otepan ot t'urskio 'ask'er*
he.was killed of turkish soldier.DIR
'he was killed by the Turkish soldier'
b. *se 'iskapif od ml'ekoto*
SELF I.spill of milk.DIR.DEF
'I spilled milk over myself'

Under the influence of other Balkan languages Balkan Slavic reconfigured the peripheral functions of the instrumental and generalized the ablative–agentive function (see Sobolev 2005: nos. 70, 76, 79, 80). Even more significant was the extension

[5] The majority of Balkan Slavic languages lost the synthetic form of the instrumental case; it is attested in several adverbs in fossilized form.

978–0–19–920647–6 50-Spencer-ch49 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 720 of 729 May 16, 2008 19:0

of the patientive function to instrumental constructions, under the influence of the general Balkan (originally Greek) model, the so-called *accusativus graecus* ('Greek accusative').

The extremes of the distribution of accusative constructions (objective, aggregate, comitative, etc. meanings of historical instrumental case) are illustrated below by examples from a Northern Greek dialect of Western Macedonia (see Sobolev 2005: nos. 46, 52, 53, 69), with Modern Russian equivalents for comparison:

(3) a. *p'etakśi tu kś'ilu* (*brosit'* *palkoj)*
threw DEF stick.ACC (throw.INF stick.INS)
'threw a stick'

b. *plir'ono lift'a* (*platit'* *den'gami*)
I.pay money.ACC (pay.INF money.INS)
'to pay with money'

c. *imbur'evome pr'ovata* (*torgovat'* *ovcami*)
I.trade sheep.ACC (trade.INF sheep.INS)
'to trade sheep'

d. *fit'evo tu xur'afi kapn'o* (*zasadit'* *pole* *tabakom*)
I.plant DEF field.ACC tobacco.ACC (plant.INF field.ACC tobacco.INS)
'to plant the field with tobacco'

e. *pirix'iɲika γ'ala* (*ja* *oblilsja* *molokom*)
I.spilled milk.ACC (I spilled milk.INS)
'I spilled milk over myself'

f. *p'otizo tis ajil'aðis nir'o* (*poit'* *korov* *vodoj*)
I.water DEF cows.ACC water.ACC (water.INF cows.ACC water.INS)
'to water the cows'

g. *ti 'ekanis ta lift'a?* (*čto* *ty* *sdelal* *s* *den'gami?*)
what you.did DEF money.ACC (what you did with money.INS)
'What did you do with the money?'

h. *ar'ostiñi varj'a ar'ostia̯* (*zabolet'* *tjaželoj* *bolezn'ju*)
became.ill serious disease.ACC (become.ill.INF serious disease.INS)
'to fall seriously ill'

In Balkan Slavic dialects this Greek-type reformation affects the semantic and functional core of the whole category, tending to oust instrumental–comitative marking from the non-core instrumental–comitative meanings or at least to anchor accusative constructions as doublets or variants of instrumental–comitative constructions. Apart from Greek dialects, the 'Greek accusative' is found in Aromanian, Albanian, Southern Bulgarian, and Macedonian dialects (Sobolev 2005: 46, 52, 53, 69, 78, 81, 82). Therefore, in certain Balkan Slavic languages, constructions which are impossible in other Slavic languages become grammatical:

(4) a. *i n'apiv g'oi̯data v'oda*
them.ACC I.made.drink cows.DIR.DEF water.DIR
'I watered the cows' (lit. 'I made the cows drink water')
b. *se k'apnaf ml'eko*
REFL I.spilled milk.DIR
'I spilled milk over myself'

The analogical spread of the instrumental case in the Balkan Slavic languages was not only restrained by Balkan influences, but even reversed. The reformation of the Slavic instrumental within the Balkan context follows the hierarchy of case meanings and affects both the core and the periphery of the corresponding semantic and functional domain. This gives rise to semantic splits and the formation of new structural oppositions not found before in Slavic languages. We observe simultaneously the effects of the patient semantic domain (accusativity),[6] following the Greek model, and the agent semantic domain (ablativity), following the Balkan Romance model. The patient/accusative domain principally affects core meanings (Macedonian and South Bulgarian dialects), while the agent/ablative domain affects peripheral meanings (Western Macedonian dialects). However, innovations are found side by side with archaisms. Even heavily Balkanized Bulgarian dialects show the old form of the instrumental in adverbial functions: *mʌn'ol səm p'ɔk'um* 'go 1SG.PRF.AUX road.INS' 'I went along the road' (Sobolev 2001: 44). Macedonian dialects highly influenced by Greek preserve instrumental–comitative encoding for objective (5) or restrictive (6a–c) meanings:

(5) *toi̯ f'ərliɬ so d'ərvo*
he threw with piece.of.wood.INS
'He threw a piece of wood'

(6) a. *se r'azbole so t'eška b'oles*
REFL became.ill with serious illness.GENERALOBL
'He/she became seriously ill'
b. *oi̯ḱ'orefne so 'očite*
we.became.blind with eyes.GENERALOBL.DEF
'We became blind' (lit. 'We became blind in (our) eyes')
c. *ž'ivei̯t so kr'adei̯ne*
lives with stealing.GENERALOBL'
'He/she lives by stealing'

In this and similar cases we can clearly observe the analogical spread of various types of verb complementation.

[6] The fields are best labelled by semantic roles (agent, patient, instrument, etc.).

49.3 Reconstructing the process of morphological case decline

A detailed analysis of variations in case marking in dialects with analytic case systems reveals innovative caseless systems (Stojkov 1968) on the one hand[7] and on the other hand, extremely rare archaic systems that preserve certain elements of morphological case in certain nouns and noun groups. From these latter systems we can trace the way innovative analytic trends develop and spread in South Slav languages, as well as the extent to which some parts of the morphological system resist the pressure towards analytic encoding. Such questions can only be addressed by considering the sum total of the instances in which inflection is preserved in these rare synthetic dialects, and by taking into consideration the competition between inflectional forms and other means of morphosyntactic marking, bearing in mind semantic and formal factors on the preservation of inflection. Unfortunately, only three to four dialects are available for this type of research, thus considerably limiting its scope.

49.3.1 Typical inflectional archaisms

The Serbian Prizren-Timok dialect of the village Svinitsa shows the following combination of inflectional archaisms of the instrumental case along with analytical innovations (the general oblique case), with morphological restrictions for certain groups of nouns:[8]

(7) a. *Is* *plug* *oremo*
with plough.M.GENERALOBL.SG we.till

b. *Is* *plug-om* *oremo*
with plough.M.INSTR.SG we.till
'We till with a plough'

(8) a. *Ga* *udril* *iz* *vil-u*
him hit with pitchfork-F.GENERALOBL.SG

[7] These systems are well-known for showing variation in their encoding of grammatical relations, as was seen above with the instrumental case example. This is due to the competition between different prepositional constructions as well as between prepositional constructions and prepositionless oblique case constructions (Sobolev 2005).

[8] The reconstruction follows Tomić (1984). This dialect (as in the majority of Balkan languages) shows neutralization of the instrumental/comitative/means case oppositions. For this reason, together with the fact that there are no examples of the instrumental in the data presented by Tomić, the examples include constructions with various functions. This does not of course affect the noun morphology.

b. *Ga udril iz vil-om*
him hit with pitchfork-F.INS.SG
'He hit him with a pitchfork'

(9) a. *Ga udril iz dzvonc-e*
him hit with bell-N.GENERALOBL.SG
'He hit him with a bell'
b. *Skočili iz nožev-e*
jumped.up with knives-M.GENERALOBL.PL
'(They) jumped up with knifes'
c. *Ona ubrisala oči iz ruk-e*
she wiped eyes with hand-F.GENERALOBL.PL
'She wiped her eyes with her hands'
d. *Ja radil iz drv-a*
I worked with firewood-N.GENERALOBL.PL
'I worked with firewood'

By looking at synthetic/analytic doublets with masculine and feminine singular nouns we can identify the following categorial restrictions on the use of oblique case: number (plural inflected forms are impossible, 9b–d) and gender (neuter nouns are indeclinable, 9a, d). In these language varieties we sometimes find that an analytic plural form is in a suppletive relationship to a corresponding synthetic singular form, for instance in the dative plural of *žena* 'wife' formed with the preposition *na* (see Table 49.1).

Table 49.1. Case paradigm of *žena* 'wife'

	SG	PL
NOM	*žena*	*žene*
ACC	*ženu*	*žene*
DAT	*ženi*	*na žene*

49.3.2 Hierarchy of morphological factors

By comparing these archaic systems we can specify those areas of nominal morphology where analytical tendencies are more (or less) conspicuous. We also gain clues as to the mechanisms and chronology of this change. Relatively archaic dialects with rudiments of a case system allow us to judge how stable given forms are in comparison with other parts of the system. From this we can formulate a typological hierarchy of morphological features that serve to maintain nominal inflection.

A comparative analysis of the noun morphology in several Serbian Prizren-Timok varieties (Sobolev 1991a) has shown that different groups of nouns (differentiated, for example, by gender and/or declension type) lose their inflection at different times, and that different stages of case loss can be found in different dialects. As a result of this investigation, we can propose the following declensional hierarchy:

a. Plural inflection is less stable than singular.
b. Inflection in neuters is less stable than that of feminine and masculine nouns.
c. Where the animacy category is intact, the inflection of inanimates is less stable than that of animates.
d. Where the person/non-person distinction is present, inflection of nouns referring to a person is the more stable.

The case stability hierarchy can be represented as follows:

NOM = ACC > DAT > GEN > INSTR = LOC.

Consideration should also be given to the stabilizing influence of formal factors. If a case that might otherwise be expected to be lost earlier has a form which is syncretic with that of some other case that is more resistant to loss, then that form may be maintained for both case functions. This is a possible explanation for the retention of the locative case forms of feminine *a*-declension nouns which are syncretic with the corresponding dative forms in the dialect of Sredska in Kosovo belonging to the Prizren-Timok dialect group.

(10) *Ide po džad-e*
goes along road-F.LOC.SG (= F.DAT.SG)
'He/She/It goes along the road'

49.4 Reconstructing the syntactic process of case decline and the hierarchy of syntactic factors

For dialects which are intermediate between synthetic and analytic (as, for example, Serbian Kosovo-Resava), it is syntactic variation and the hierarchy of syntactic contexts affecting the preservation of semantic role marking in nominal inflection that provides the best heuristic clues as to how the systems are organized. Case functions in Slavic need to be analysed against the backdrop of all those competing synonymous constructions, which, in theory at least, may be found side by side in that variety. Thus, in analysing the possessive genitive (*sin babe* 'son old.woman.GEN'

'the old woman's son') it is necessary to consider all the other possible ways of encoding possessive relationships in the noun phrase, such as possessive adjectives (*babin sin* 'old.woman.POSS.ADJ son'), constructions with preposition *od* + genitive (*sin od babe* 'son of old.woman.GEN'), constructions with preposition *od* + general oblique case (*sin od babu* 'son of old.woman.GENERALOBL'), the possessive dative (*sin babi* 'son old.woman.DAT'), and possessive constructions with the preposition *na* + general oblique case (*sin na babu* 'son on old.woman.GENERALOBL'), which may be furnished with a shortened form of a personal pronoun (*sin i na babu* 'son her.DAT on old.woman.GENERALOBL').

The distribution of morphological, morphosyntactic, and syntactic items in a given variety has to be supplemented with the complete set of semantic, formal, lexical, communicative, etc. restrictions on use.[9] Empirical work in this domain is rendered very difficult by the time-consuming nature of the data collection techniques (see Belyavski-Frank 1983; Sobolev 1991a; Miloradović and Greenberg 2001), which includes obtaining negative information (i.e. those forms and constructions which are *not* used in a given variety), as well as the problem of filtering out items which have entered into that variety from a literary standard, and the identification of fossilized and adverbialized forms.

Comparison of these varieties allows us to establish the place of each case form and preposition + case construction on the hierarchy, and to determine its stability compared with its analytical counterpart. Such a hierarchy has been proposed for the Southern Slavic genitive, locative, and instrumental cases in Sobolev (1991a):

a. Genitive: the most stable is the partitive genitive with a noun; less stable is the adnominal genitive object and the prepositional construction.
b. Instrumental: the most stable are the prepositionless prosecutive (cf. 1i) and the temporal (cf. 1j),[10] less stable is the instrumental–comitative construction *s* + instrumental.
c. Locative: three prepositional constructions (*o*, *po*, *pri* + locative) which mark syntactic roles by both inflection and by preposition are extremely stable.

A typical set of forms, with their distribution and their usage in prepositional constructions for transitional analytic/synthetic Serbian Kosovo-Resava dialects is found in Miloradović (2003). The usage of prepositional and prepositionless forms of the genitive are polarized: synthetic forms are found predominantly in prepositional constructions (therefore the syntactic function is marked both by inflection and the preposition), while the tendency is to use the innovated analytic form for the prepositionless construction.

[9] For example, use of the attribute 'old' in *sin stare babe* makes the usage of a possessive adjective impossible, **stara-babin sin*.

[10] The instrumental in these functions has a tendency towards adverbialization in Slavic.

The construction with preposition *na* 'on' + general oblique case instead of the synthetic dative is only occasionally attested as the realization of the benefactive argument of a verb, and then only when it co-occurs with distributive semantics.[11]

(11) *Davam tam na decu ja, na đaci*
I.give there on children.GENERALOBL I on pupils.GENERALOBL
slatkiši
sweets.GENERALOBL
'There, I give sweets to the children, to the pupils'

It is indicative that this construction is never used with speech act verbs (*verbi dicendi*), which allow only the dative case (this applies to plural arguments, too):

(12) *javite Boškovićima*
tell.IMP Bošković.DAT.PL
'Tell it to (the family of) Boškovićs'

The instrumental case without preposition is still used in this variety in the instrumental meaning, where it competes with the constructions *s* + instrumental and *s* + GeneralObl, in the prosecutive meaning, *idu putem* 'go.3PL road.INS' 'They go along the road' and in the objective meaning, *klima rukama* 'wave.INF hands.INS' 'to wave one's arms' (lit. 'to wave with arms').

When we look at the formation of the general oblique case as a syntactic device for ousting synthetic forms in favour of analytic constructions it is clear that this process does not lead to the emergence of an analytical structure. It only limits the domains in which oblique cases are used and shifts them into the areas where analytic constructions cannot penetrate. However, it does not eliminate them as morphological categories. An interesting example is a complete elimination of the synthetic instrumental form by the general oblique case following the prepositions *pod* 'under' and so on (13a, 14a, b) as compared to the exclusive use of the instrumental case with the preposition *za* in constructions denoting motion following something or somebody (13b):

(13) a. *pod ljuljku zmija bila*
under cradle.GENERALOBL snake was
'There was a snake under the cradle'
b. *pođe za ocem*
went after father.INS
'He went after, behind his father'

[11] The same shade of meaning is attested in the similar prepositional constructions which compete with the synthetic benefactive dative–genitive in Modern Greek and Aromanian (Sobolev 2005).

(14) a. *pođeš za devojku*
you.go after girl.GENERALOBL
'You (will) marry a girl', lit. 'You (will) go after a girl'
b. *iem za krompiri*
I.go after potatoes.GENERALOBL
'I go for potatoes'

This specialization is due to the need to differentiate the spatial *za*-construction from the objectival construction, which permits only the general oblique case.

Typically, the locative is absent following the prepositions *u* 'in' and *na* 'on'. Following the preposition *po* 'in, across', the locative and general oblique case are interchangeable:

(15) a. *tamo ima po sel-u*
there have in village-LOC
'There is in the village . . . '
b. *zbiru jaja po selo*
collect eggs in village.GENERALOBL
'They collect eggs in the village'

Only the locative is found after the preposition *pri* 'at', never the general oblique case.

Of the cases that are used with prepositions the genitive occupies a special position, in that it is always possible to relace the genitive with the general oblique. For other cases in prepositional uses (e.g. instrumental and locative) there will always be at least one syntactic environment in which the alternation with the general oblique is not possible, for example:

(16) a. *potrčim za kolima*
I.run after car.INS
'I run after, behind the car'
b. *pri guši nosi lančić*
at neck.LOC s/he.wears chain
'He/she wears a little chain on his/her neck'

49.5 Typological verification of historical evidence

By observing archaic synthetic constructions in the modern analytic dialects (such as Serbian Prizren-Timok) and by observing the way dialects such as Serbian

Kosovo-Resava develop analytic constructions we can verify the individual data in medieval South Slavic (Old Serbian, Old Bulgarian, Old Macedonian) written sources. We can also reconstruct language change in modern analytic South Slavic languages, and we can determine the systemic significance of specific pieces of evidence taken from the primary historical sources. In other words, the dialect data can be used for verifying or even reconstructing the loss of noun inflection in Balkan Slavic as we know it from the medieval sources. For example, the usage of a form in the general oblique case following preposition *pri* in the Bojana inscription of 1258 (Popkonstantinov and Kronsteiner 1997) clearly shows that the analytic construction had infiltrated the most stable domain of locative usage (see section 49.4, 'c') and from this we can conclude that analytic processes were occurring in other domains, too, notwithstanding the lack of direct evidence:

(17) *napisa že sę pri c[a]rьstvo*
wrote EMPH REFL at kingdom.GENERALOBL
blgarskoe[12]
Bulgarian.ADJ.GENERALOBL
'and it was written in the time of the Bulgarian kingdom'

Likewise, we can demonstrate that analytical constructions were expanding in Bulgarian and Macedonian by the twelfth–thirteenth centuries. Examples (18) and (19) show prepositions followed by the general oblique case instead of, respectively, the genitive and instrumental:

(18) *ispousti golǫbě i̯s kovčegъ*
released dove.ACC from ark.GENERALOBL
'He released the dove from the ark'

(19) *ikonǫ sъ obrazъ pokazaǫšte*
icon.ACC with sacred.image.GENERALOBL showing
'Showing the icon with the sacred image'

There are at least two ways of interpreting the synchronic emergence of morphological restrictions on the syntactic realization of given structures. The first lies in the noun morphology: we may have neutralization of the archaic oblique case and the general oblique, that is, between analytic and synthetic encoding strategies. Another possibility might be a gradual increase in analytic constructions in a given variety, perhaps due to such factors as analogy or calquing and the resulting pressure on the morphological level in that variety. However, the only factor that leads to a complete loss of inflection is the phonetic identity of the oblique case

[12] Instead of *pri carьstvě blьgarьscěmь* 'at kingdom.LOC Bulgarian.LOC'

978–0–19–920647–6 50-Spencer-ch49 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 729 of 729 May 16, 2008 19:0

with the accusative case within a certain group of nouns, which is later reanalysed as a form of the general oblique case. The key feature in the history of declension in Bulgarian and Macedonian is that the course of attrition, which we know to be 'natural' from the evidence of Serbian dialects in which certain noun groups undergo this loss in turns, was speeded up by the fact that the innovation began to spread from the most conservative and linguistically stable groups of nouns, such as masculine singular animates and feminine nouns in the so-called *a*-declension (Sobolev 1991b).

CHAPTER 50

CASE IN AN AFRICAN LANGUAGE

IK – HOW DEFECTIVE A CASE CAN BE

CHRISTA KÖNIG

50.1 Introductory remarks

Africa is known for being a continent where there are not so many languages with a grammaticalized case system. In East Africa there is one language which is quite exceptional. Not only does it have a case system distinguishing seven cases, which on African standards is a lot, but it is also unusual in that nearly all elements in the language are case-inflected, including nouns, verbs, adverbs, prepositions, and conjunctions. Nevertheless even though case is ubiquitous in Ik, the encoding of core participants, that is intransitive subjects (S), transitive subjects (A), and transitive objects (O), is so defective that it is questionable whether the elements under consideration are really cases and not something else.

Ik is a Kuliak language[1] spoken in northeastern Uganda in the border area between Uganda, Kenya, and Sudan. Its exact genetic status is uncertain. According

[1] The term Kuliak was invented by Heine (1976).

to Greenberg (1963a), it belongs to the Eastern Sudanic branch of the Nilo-Saharan phylum, while Laughlin (1975) proposed to leave Kuliak unclassified. Ik has a basic verb-initial constituent order (AVO/SV). It has a rich system of verbal derivation, and it has voiceless vowels (which will be represented by raised characters, e.g. ^{e}). These vowels are crucial in the case system. All elements in the language are realized in two forms, one called final form the other one non-final form. In principle, the final form occurs before pauses, for example at the end of clauses or noun phrases; in all other environments the non-final form occurs. The following treatment is based on König (2002) where further details are outlined.

50.2 Case

Ik has an elaborate case system. Table 50.1 gives an overview of the seven case suffixes distinguished.

Table 50.1. The case inflections of Ik

CASE	Abbreviation	Final	Non-final
Nominative	NOM	*V-$^{[a]}$	*V-[a]
Accusative	ACC	**-k**$^{[a]}$	[-a]
Dative	DAT	**-k**^{e}	**-e**
Genitive	GEN	**-e** (-i)	**-e**
Ablative	ABL	**-o** (-u)	**-o**
Copulative	COP	**-k**^{o}	**-o**
Oblique	OBL	ø	ø

Each suffix occurs in two forms: the final form and the non-final form. Basically, Ik is an accusative language, that is, S and A are treated similar and simultaneously different from O. This means for main clauses that the nominative encodes A (cf. example (1)) and S (2), while the accusative encodes O (1). However, there are many contexts where the accusative pattern is neutralized: If the subject, S or A, refers to the first or second person, all core participants (S, A, and O) occur in the nominative (see (3)–(7)). I have used the term 'case anomaly' (König 2002) for this neutralization of case. In imperative and cohortative clauses with VAO/VS order, all core participants, S, A, and O, are encoded in the oblique case (see (8) and (9)). The oblique case is the only non-derived case form in Ik (see Table 50.1). It has also been called the basic form of the noun by Heine (1983) and Serzisko (1992). Often the noun shows a final vowel which in other case forms is deleted. A and S are usually omitted in imperatives, but they can be expressed. In imperative and cohortative clauses with AVO/SV order, A and S occur in the nominative and O in the oblique

978–0–19–920647–6 51-Spencer-ch50 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 732 of 741 May 16, 2008 19:0

(see 10). In relative clauses and other subordinate clauses marked by the subjunctive expressed by a verbal suffix *-ik*^{e} indicating subordination, all core participants (S, A, and O) are encoded in the accusative (for relative clause see O = ACC in (11a) and (11b); A = ACC in (11b); S = ACC in (12); for subjunctive clause see O = ACC in (13) and (14); A = ACC in (14); S = ACC in (15)). In clauses with a topicalized object, all core participants occur in the nominative (see 16). Topicalized participants occur preverbally in the nominative (see 18). Examples (17) and (18) convey the same content: (17) in the default order, and (18) with a topicalized object which is left dislocated and appears in the nominative. The conjunctions of the language trigger different case patterns, depending on the verb form, that is, whether the verb occurs in the subjunctive or the narrative or the optative. Complement clauses and auxiliary clauses show the same kind of complexity.

Main clause

(1) *en-es-uɠot-a wík-á njíní-k*^{a}.
see-IRR-AND-a children-NOM we.INCL-ACC
'The children will see us (incl).'

(2) *Mɛt-és-íd-a bi-a.*
be.ill-IRR-2SG-a you-NOM
'You (sg) will be ill.'

(3) *en-és-isín-a njín-*^{a} *wík-*^{a}
see-IRR-1PL.INCL-a we.INCL-NOM children-NOM
'We (incl.) will see the children.'

(4) *en-és-im-a ngw-*^{a} *wík-*^{a} *awá-ɔ.*
see-IRR-1PL.EXCL-a we.EXCL-NOM children-NOM home-ABL
'We (excl.) will see the children at home.'

(5) *en-í-a nk-*^{a} *wík-*^{a}
see-1SG-a I-NOM children-NOM
'I see the children.'

(6) *en-es-íd-a bi-*^{a} *wík-*^{a}
see-IRR-2SG-a you-NOM children-NOM
'You (sg.) will see the children.'

(7) *en-és-ít-a bit-*^{a} *wík-*^{a}
see-IRR-2PL-a you(pl)-NOM children-NOM
'You (pl.) will see the children.'

Imperative or cohortative clause

(8) *ats-é bi.*
come-IMP.2SG you.OBL
'(You) come!'

978–0–19–920647–6 51-Spencer-ch50 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 733 of 741 May 16, 2008 19:0

(9) *en-é bi wíce.*
see-IMP.2SG you.OBL children.OBL
'(You) see the children!'

(10) *bi-á ɠa-ée saɓ-ée loŋót*^{a}.
you-NOM go-IMP.2SG kill-IMP.2SG enemies.OBL
'(You) go and kill enemies!'

Relative clause

(11) a. *cek-a ná ntsí wícé-á en-uɠot-í bíra-a*
woman-NOM REL.SG she.OBL children-ACC see-AND-1SG be.not-a
nɛɛ́ na.
here.DAT DEM
'The woman whose children I saw is not here.'
b. *cek-a ná ncí-a en-uɠot-í-á ntsí*
woman-NOM REL. SG I-ACC see-AND-1SG-a she.OBL
wice-k^{a} *bíra -a nɛɛ́ na.*
children-ACC be.not-a here.DAT DEM
'The woman whose children I saw is not here.'

(12) *ɠan-uɠɔ́t-io nyɔt-a nú ekw-itíní-a maraŋ-ak-á*
take-AND-NARR men-NOM REL.PAST eye-PL-ACC good-PL-a
nyɛr-a ní léɓetse
girls-ACC REL.PL two.OBL
'And the men whose eyes were good took the two girls.'

Subjunctive clause

(13) *ɠ-á cɛ́k-*^{a} *en-íe wicé-k*^{a}.
go-a woman-NOM see-SBJV children-ACC
The woman goes when she sees the children.'

(14) *na nci-a en-í-ik*^{e} *wicé-k*^{a} *ɠo-i-ak*^{o}
when I-ACC see-1SG-SBJV children-ACC go-1SG-NARR
'When I see the children I go.'

(15) *na wicé-á ni ats-át-ik*^{e} *kɔ́ŋ-ɛsɛ tɔ́bɔŋ-a*
when children-ACC DEM come-3PL-SBJV cook-NARR.IPS food-NOM
ńtí-k^{e}.
they-DAT
'When these children came food was cooked for them.'

Topic clause

(16) *wík-a ńc-i en-a ná ńts-*^{a}.
children-NOM I-GEN see-a ENC he-NOM
'As for my children, he sees (them).'

Clause type A: Focus clause, relative clause, clause with the subjunctive, clause with the dummy pronoun triggered by the conjunction, imperative or cohortative clause, object topic clause.

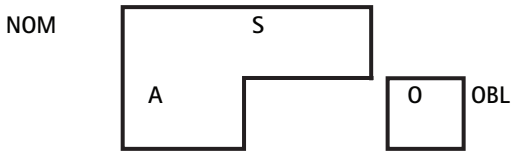
I Focus clause, relative clause, clause with a dummy pronoun triggered by the conjunction:
S = A = O



II Imperative and cohortative clause in a VA/S-order and *alaké*-clause with optative:
S = A = O



III Imperative and cohortative clause in an A/SV-order, an optional variant of II:
S = A
O ≠ S, A

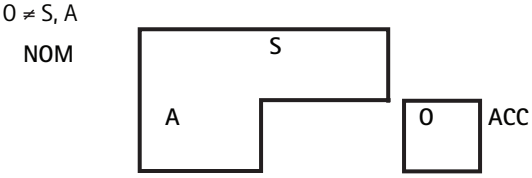


IV Object topic clause:
S = O = A



Clause type B = contexts other than **A:**

V for S, A = third person
S = A
O ≠ S, A



VI for S, A ≠ third person (= **case anomaly**):
S = A = O

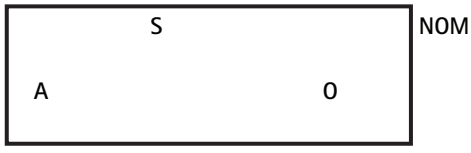


Figure 50.1. Case patterns in Ik

978–0–19–920647–6 51-Spencer-ch50 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 735 of 741 May 16, 2008 19:0

(17) *ŋʊ̀r-í-a na ńts-^{a}.*
break-1SG-a ENC he-NOM
'I cut him [the hair].' (Lit. 'I broke him.')

(18) *ńtsa-á ŋʊ̀r-í-á nak^{a}.*
he-NOM break-1SG-a ENC
'Him I cut [the hair].' (Lit. 'He, I broke.')

Generally speaking, the coding of core participants is triggered by various factors: the person-marking properties of the subject, constituent order, TAM-marking on the verb, whether one is dealing with subjunctive, narrative, optative mood, or clause type. In total, five different case patterns are used to encode S, A, and O; they are illustrated in Figure 50.1. The rules apply in a hierarchical order. First, it is of importance whether the clause belongs to type A or B. Unfortunately type A is not uniform, therefore each of its features has to be listed separately. Types A and B are mutually exclusive. Type B contains main clauses but also subordinate clauses, for example when the verb is used in the narrative. Only two of the five patterns are case-sensitive, namely patterns III and V. In all other patterns, case is neutralized as A, S, and O are always treated identically. With regard to the cases which occur to encode the core participants the following are used: the accusative, the nominative, and the oblique.

50.3 Alternative analysis

One may wonder whether Ik is a case language if there are so many contexts in which case is neutralized. There are the following reasons for maintaining a case analysis: First, there is no alternative which would be more adequate than one in terms of case. There have been attempts, in particular by Serzisko (1992), to interpret the nominative and the accusative as discourse-pragmatic markers. The function of the accusative is only vaguely described by him, the nominative is a '*diskurspragmatischer Marker des präsentierten Partizipanten in thetischen Aussagen*'.[2] As has been shown in König (2002), narrative discourse data do not support Serzisko's analysis. Table 50.2 presents the frequency of the cases used to encode S, A, and O in one particular narrative discourse. The results corroborate the case analysis: in narrative discourse, the nominative is the default case to encode S and A. The accusative is the default case to encode O.

[2] The fact that the ACC is no accusative in terms of case of an object results from the fact that it can mark as well the *actor* as the *undergoer.* We do not know its exact function and won't currently give attention to it.' (Serzisko 1992: 172)

978–0–19–920647–6 51-Spencer-ch50 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 736 of 741 May 16, 2008 19:0

Table 50.2. Encoding of S, A, and O in the narrative text 'The three girls'

Case	S	A	O	N.PRED	Other
NOM	28	16	3		4
ACC	5	2	35		
OBL	3 *ńda*		3	1	16
GEN			2	1	
COP				5	3

With regard to frequency, irregularities are much less common than the five patterns would suggest. Second, taking the whole case paradigm into account (as presented in Table 50.1), peripheral participants behave very regularly. Third, even if the core cases in particular are defective, there are obligatory syntactic rules which determine their occurrence. In this respect they fulfil the case definition.[3] Fourth, from a typological perspective, in Africa most accusative languages are les homogeneous than a Eurocentric perspective might suggest. By African standards it is more typical than not to have a split system with neutralized contexts (see König forthcoming, a). In sum, case seems to be the best option to describe the system under consideration. It goes without saying that, in addition to its syntactic function, case also has a pragmatic value (see section 50.4).

As in other languages, case in Ik is a category which is connected with nouns and pronouns, but case is not restricted to these word classes. Function words like conjunctions, postpositions, prepositions, adverbs, and even verbs are also inflected for case. Examples (19) and (21) illustrate the manifestation of case with regard to the case-inflected conjunction *toimɛn* 'that'. *Toimɛn* occurs in the case required by the syntax, namely nominative in (19), accusative in (20), and dative in (21).

(19) ***toimɛn-a*** *ńtá* *tóped-uʤo-íd-i* *ɪt-ɛ́s-*[a].
that-NOM NEG be.able-AND-2SG-NEG reach-INF-NOM
'(that) you are not able to reach [there].'

(20) *ɪtámáán-á* *mo* *tam-i* ***tóimɛní-k***[a] *ińta* *en-í-í* *nts*[a].
must-a NEG think-NEG that-ACC NEG see-1SG-NEG he.NOM
kɔn-ɛn-ɔ *á-i.*
one-PEE.SG-COP side-GEN
'He need not think that I will not find him anywhere.'

[3] Case is an inflectional **system** of marking nouns/noun phrases for the type of relationship they bear to their heads. Inflectional systems are expressed by affixes, tone, accent shift, or root reduction; adpositional systems are included only if they encode core participants such as S, A, and O. (König 2008a).

(21) *ɪtét-í-a ńa* ***tóimɛní -k^{e}*** *ńʤ-a nyéʤa bi-k^{a}*.
notice-1SG-a ENC.SG **that-DAT** eat-a hunger-NOM you-ACC
'I noticed that you felt hungry' (Lit: 'hunger ate you).'

Table 50.3 gives an overview of case-inflected elements in Ik. The left column presents the ungrammaticalized source concept, a noun or a case marker; the middle column presents the grammaticalized function of the source concept, and the right column presents the case inflections in which the grammaticalized function occurs. Two cases, the dative and the copulative, have become part of verbal inflection: the dative has been grammaticalized to a subjunctive marker (as in (13) to (15)) and the copulative has given rise to the narrative.

Peripheral participants are marked throughout by case. The ablative and dative are both used in a wide range of different functions (up to twelve different

Table 50.3. Case inflected items in Ik

Source			Target		Case inflections accepted
Dative	*-k^{e}/-e*		Subjunctive	*-ik^{e} -ie*	
Copulative	*-k^{o}/-o*		Narrative	*-uo*	
Noun			**Case-inflected conjunction**		
	tʊ́mɛd a (na)	?		'where'	NOM, ACC
	mɛná	'thing'		'what'	NOM, ACC
	kɔrɔɓáa	'thing'		'what'	NOM, ACC
	na	'place'		'where'	NOM, ACC, DAT, ABL
	tóimɛn	'problem'		'that'	NOM, ACC, DAT
Noun			**Case-inflected adverb**		
	wash	'front'		'ahead', 'first', 'earlier'	DAT, ABL, COP, GEN
	na	'place'		'here'	NOM, ACC, DAT, ABL, COP, OBL
	yas^{i}	'truth'		'true', 'really'	DAT, COP
	[nominal source no longer known]		*ɛɗá*	'alone'	COP, GEN
			muny^{u}	'all', 'completely'	OBL, COP
			ɟík^{e}	'always'	DAT, GEN
			kooke	'there'	DAT invariable
Relational Noun			**Case-inflected postposition/preposition**		
	ai^{a}	'side'		'from'	all
	aƙw	'palm (of hand)', 'sole'		'inside'	all
	búbú	'stomach'		'under'	all
	ɠwarí	'surface'		'top'	all
	kan	'back'		'behind'	all

978–0–19–920647–6 51-Spencer-ch50 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 738 of 741 May 16, 2008 19:0

functions) (see Heine 1990). I will illustrate the range with regard to the ablative. The ablative encodes source (see (22)), the sender in (23), the locative in (24), the instrument in (25), the partitive in (26), the cause in (27), manner in (28), time in (29), the agent in passive clauses in (30), the possessor in verbal possession in (31), and the standard in comparative expressions in (32). As in other case languages as well, the verb determines which role is expressed by which case. Location for instance is encoded with some verbs in the dative ((33)) with others in the ablative, as in (24).

(22) *ŋata na kan-ed-o aw-è njín-i.*
run ENC back-ABL compound-GEN we.EXCL-GEN
'He runs back to our compound.'

(23) *dzígw-í-a na híɔ̄^{a} túde wice ai-u.*
transfer-I-a ENC cows-NOM five children.OBL side-ABL
'I buy five cows from the children.'

(24) *ép-a nts-á ho.*
sleep-a he-NOM house.ABL
'He sleeps in the house.'

(25) *ngʃagʃ-a ńts-a tɔbɔŋw-á golom-^{o}.*
eat-a he-NOM food-ACC wooden.spoon-ABL
'He eats food with a wooden spoon.'

(26) *ma-i-á bi-e lɔtɔ́ba-o.*
give-1SG-a you-DAT tobacco-ABL
'I gave you some [of the] tobacco.'

(27) *bad-uʛot-á nyéʛe-o.*
die-AND-a hunger-ABL
'He died of hunger.'

(28) *íŋarés-et-a be ncí-e lakás-on-^{o}*
help-3PL-a ENC I-DAT fun-INF-ABL
'They helped me with pleasure.'

(29) *barats-o nak^{a} nyaɓaít-^{o} ʛo-í-a kakum-e ɛdá.*
morning-ABL ENC dawn-ABL go-1SG-a Kakuma-DAT alone
bíra-e íŋaresí-k^{a}
be.NEG-DAT help-ACC
'Early in the morning I went alone to Kakuma without any help.'

(30) *wa-ós-a dakwa ná nc-u.*
harvest-PASS-a tree.ABS ENC I-ABL
'The tree is harvested by me.'

978–0–19–920647–6 51-Spencer-ch50 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 739 of 741 May 16, 2008 19:0

(31) *bíra dakw-a nc-u.*
be.NEG tree-NOM I-ABL
'I have no tree.'

(32) *mɪt-a nyarama na da njin-ú.*
be-a girl.OBL REL be.beautiful we.INCL-ABL
'She is the most beautiful of us.'

(33) *i-á hɔ-k^{e}*
be-a house-DAT
'He is in the house.'

50.4 Other means

There seems to be a tendency that case comes into play if other means fail to distinguish S, A, and O. Other means are in particular cross-referencing of S and A. The verb allows cross-referencing of the subject (S and A) only. For subject (S and A), being first or second person, cross-referencing is obligatory. Third person singular cannot be cross-reference. Third person plural can be cross-referenced, but there are complex rules which determine the occurrence of the third person plural bound pronoun. Table 50.4 gives an overview of the suffixes which are used to cross-reference S or A on the verb.

Table 50.4. Bound pronouns in Ik

		Final	Non-final
SG	1	-í	-íá
	2	-íd	-ídà
	3	Ø	-à
PL	1INCL	-ísín	-ísínà
	1EXCL	-ím	-ímá
	2	-ít	-ítá
	3	-át	-átà

In main clauses, the third person plural suffix *-át^{a}* is ungrammatical in a VA/VS-order (compare (34d) and (34e)). It is optional in an SV/AV-order if expressed nominally (35b). In relative clauses it is optional too (36a, b). It is obligatory if not expressed by an independent noun (34a), if expressed by a pronominal subject

978–0–19–920647–6 51-Spencer-ch50 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 740 of 741 May 16, 2008 19:0

Table 50.5. Continuum of the occurrence of the third person plural suffix *-át*^{a} in Ik

I Ungrammatical	II Optional	III Obligatory
Nominal subject VS/A	a. Nominal subject S/AV b. REL-clause	a. No independent subject b. Pronominal subject c. Complement clause Comparative Subjunctive

(compare 34b, c), as well as generally in the following clause types: comparative (37), subjunctive (34c), and complement clauses. As Table 50.5 illustrates, the rules for the occurrence of the third person plural bound pronoun range from being ungrammatical to being optional up to being obligatory.

Main clause

(34) a. *en-es-át-a ceki-k*^{a}
see-IRR-3PL-a woman-ACC
'They will see the woman.'

b. *en-es-át-a iɲt-*^{a} *ceki-k*^{a}
see-IRR-3PL-a they-NOM woman-ACC
'They will see the woman.'

c. **en-es-a ɲt-*^{a} *ceki-k*^{a}
see-IRR-a they-NOM woman-ACC

d. *en-es-a wik-a ceki-k*^{a}
see-IRR-a children-NOM woman-ACC
'The children will see the woman.'

e. **en-es-át-a wik-a ceki-k*^{a}
see-IRR-a children-NOM woman-ACC

(35) a. *mɪt-á kúrúɓá-a ní ntí*
be-a things-NOM DEM they-GEN
'These things belong to them.'

b. *kúrúɓá-a ní mɪt-át-a ntí.*
things-NOM DEM be3PL-a they.GEN
'These things belong to them.'

(36) a. *roɓ-a ní dun-et-*^{a}
people-NOM REL.PL old-VEN-a
'People who grow old . . . '

b. *roɓ-a ní dun-et-at-*^{a}
people-NOM REL.PL old-VEN-3PL-a
'People who grow old . . . '

978–0–19–920647–6 51-Spencer-ch50 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 741 of 741 May 16, 2008 19:0

(37) *ŋɪɬ-ak-á* *nyɔt-a* *ɪlɔ-át-a* *cɪk-ámá-k*^{a}.
be.strong-PL-a men-NOM defeat-3PL-a woman-PL-ACC
'Men are stronger than women.' (Heine 1999: 44)

Taking case and cross-referencing mechanisms into account, the hypothesis proposed above is sustained: if a core participant is encoded via cross-reference, case is defective. This holds in particular for first and second person. If cross-reference is defective, case encodes the core participants. This holds for third person, especially third person singular, but also third person plural. With the latter, in particular whenever case fails, *-át*^{a} is obligatory. As has been shown above, subjunctive and relative clauses are defective with regard to case. If there is no subject expressed by an independent element, use of *-át*^{a} is obligatory. With regard to cross-reference, Ik also shows an accusative pattern since the only participants being cross-referenced are S and A.

50.5 Conclusions

Ik has basically a (nominative/)accusative system which shows some unusual features. There are many syntactic contexts in which the basic case opposition between the nominative and the accusative is neutralized. Therefore, more precisely speaking, Ik is a split-accusative language, meaning that it either shows an accusative pattern or no distinction at all. Split conditions occur in particular with respect to person. In main clauses, the accusative system is only present if the subject is a non-participant, that is, if it is neither first nor second person. Clause type, TAM-marking, the presence of the dummy pronoun, and constituent order are further conditioners for split-accusativity. Five different case patterns are used to encode the core participants S, A, and O. In three of them, the case distinction is neutralized. Case is a highly productive mechanism of Ik. Nearly all lexical items of the language can be case- inflected, including adverbs, conjunctions, adpositions, and verbs.

CHAPTER 51

DIFFERENTIAL CASE-MARKING OF ARGUMENTS IN AMHARIC

MENGISTU AMBERBER

51.1 Introduction

In this chapter we examine the case and agreement systems of Amharic with particular reference to the phenomena of differential subject marking (DSM) and differential object marking (DOM). We focus on three interrelated issues: the semantic factors that may be relevant to DSM and DOM, the interaction between DSM/DOM and agreement on the verb, and the interaction between DSM/DOM and topicality will be examined.

The chapter is organized as follows. In section 51.2 a brief typological and genetic background of the Amharic language is presented. In section 51.3, the presence of the grammatical relations subject and direct object will be established on the basis of language-internal formal evidence. In sections 51.4 and 51.5 respectively differential object marking and differential subject marking will be investigated. In section 51.6, we conclude with a brief summary of the main points of the chapter.

978–0–19–920647–6 52-Spencer-ch51 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 743 of 755 May 16, 2008 19:0

51.2 Typological background

In this section, a brief typological overview of the Amharic language is presented in order to provide the reader with basic information about the grammar of Amharic that is relevant to the present topic. For a more detailed description of the language the reader is referred to Leslau (1995).

Amharic is a nominative/accusative language and exhibits a mixture of dependent- and head-marking. A definite object NP is obligatorily marked by the accusative suffix *-n*. The subject NP is unmarked. As can be seen in (1), the verb obligatorily takes subject pronominal affixes which cross-reference the subject NP. Object/indirect object marking on the verb is often optional.

(1) a. *aster wədə bet hed - əčč*
Aster to home go.PRF-3F
'Aster went home'
b. *aster bet-u-n ayyə-čč-(iw)*
Aster house-DEF-ACC see.PRF-3F-3MO
'Aster saw the house'

The morphological structure for a main clause verb in the perfective (past) conjugation is: [verb stem + subject + (object)] as shown in (2):

(2) *səbbər-əčč-(iw)*
break.PRF-3F-3MO
'She broke (it)'

The unmarked constituent order in a simple declarative clause is SOV (subject–object–verb). NP modifiers, such as possessor NPs, adjectives, demonstratives, numerals, the relative clause, precede the head noun.

Amharic has several positional prefixes (adpositions). Some of the basic ones are exemplified below (see also Hetzron 1970):

(3) *lə-* 'to' or 'for':
a. Recipient: ***lə-****liǰ-u* *sət'ət'ə-hu-t*
to-man-DEF give.PRF-1SG-3MO
'I gave it to the man'
b. Beneficiary: ***lə-****liǰ-u* *sət'ət'ə-hu-llə-t*
for-boy-DEF.M give.PRF-1SG-for-3MO
'I gave it (to someone) for the (benefit of) the child'

(4) *bə-* has a number of meanings including 'with, by means of, by, through, in/at, into':
a. Instrumental: ***bə*** - *bilawa* 'with a knife'
b. Location of an event: ***bə*** - *gəbəya* 'at a market'

978–0–19–920647–6 52-Spencer-ch51 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 744 of 755 May 16, 2008 19:0

(5) *kə-* 'from, origin of place, out of' (source):
kə-*bet mət't'a*
from-house come.PRF.3M
'He came from the house'

(6) *wədə* indicates direction away from speaker (allative):
wədə *tərar-očč-u* *hed-ə*
to mountain-PL-DEF go.PRF-3M
'He went to the mountains'

There are also a number of adpositions which are composed of a prepositional prefix and a postpositional element (of nominal origin) as in (7):

(7) ***bə****-t'ərəpp'eza-w* ***lay***
on-table-DEF **on**
'On the table'

In constructions such as (7) often the prepositional element can be omitted without affecting the adpositional meaning of the phrase.

51.3 Grammatical relations: coding strategies

51.3.1 Subjects in Amharic

It is generally assumed that the subject is a grammatical relation that expresses the A and S grammatical functions (Blake 2001; Dixon 1994; Andrews 1985a, 2007).

In Amharic, A of transitive clauses and S of intransitives can be distinguished by the three coding strategies: clause-initial position, nominative case, and obligatory verb-agreement:

(8) a. *ləmma gəməd-u-n k'orrət' -ə*
Lemma rope-DEF-ACC cut.PRF-3M
'Lemma cut the rope'
b. *ləmma mət't'a*
Lemma come.PRF.3M
'Lemma came'

In both (8a) and (8b), the NP *ləmma* occurs in clause-initial position. It gets the unmarked nominative case and controls the choice of pronominal marker on the verb. When the verb is passivized, the patient NP becomes the subject and shows

the same coding features as an 'agentive' subject: it occurs clause-initially, has nominative case, and agrees with the verb as in (9):

(9) *gəməd-u tə-k'orrət'-ə*
rope-DEF PASS-cut.PRF-3M
'The rope was cut'

The coding strategies confirm that there is a grammatical relation 'subject' in Amharic that cannot be reduced to semantic factors.

51.3.2 The Direct Object in Amharic

In terms of the three coding strategies – word order, case, and agreement – it can be established that there is a grammatical relation 'direct object' in Amharic.

In Amharic, the direct object of a transitive verb such as 'break' normally occurs after the subject and before the verb.

(10) a. *ləmma and t'ərmus səbbər -ə*
Lemma one bottle break.PRF-3M
'Lemma broke one bottle'
b. *ləmma t'ərmus-u-n səbbər -ə*
Lemma bottle-DEF-ACC break.PRF-3M
'Lemma broke the bottle'

Notice that in (10a) there is no morphological distinction between the subject and the direct object as neither is overtly marked for case. However, in (10b) we see that the object receives the object marker ('accusative' case) *-n* when it is definite. The relevant parameter is 'definiteness' rather than 'animacy', as can be seen in the following example:

(11) *ləmma and lɨǰ-(*ɨn) ayy-ə*
Lemma one child-(ACC) see.PRF-3M
'Lemma saw one child'

In (11) the direct object is animate but indefinite. It cannot be marked by the accusative case. Notice also that while there is a subject agreement marker on the verb, there is no corresponding object marker. However, the verb is optionally marked for object agreement provided that the direct object is definite:

(12) a. *ləmma t'ərmus-u-n səbbər-ə-(w)*
Lemma bottle-DEF-ACC break.PRF-3M-3MO
'Lemma broke the bottle'
b. *ləmma and t'ərmus səbbər-ə-(*w)*
Lemma one bottle break.PRF-3M-(3MO)
'Lemma broke one bottle'

978–0–19–920647–6 52-Spencer-ch51 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 746 of 755 May 16, 2008 19:0

Thus, when the direct object is not definite as in (12b) there is no object agreement marker on the verb. However, the object agreement marker on the verb becomes obligatory when the direct object is not overtly expressed:

(13) *ləmma səbbər-ə-*(w)*
Lemma break.PRF-3M-3MO
'Lemma broke it'

Thus, in constructions such as (13), the object marker is not merely an expression of agreement morphology but rather itself functions as a pronominal – an argument of the verb. When the verb is passivized the direct object becomes the subject and controls subject agreement as expected:

(14) *t'ərmus-u tə-səbbər-ə*
bottle-DEF PASS-break.PRF-3M
'The bottle broke'

In the following section we explore the formal and semantic properties that may dictate the presence or absence of the accusative suffix and the interaction between the three coding strategies – word order, case, and agreement with respect to the direct object grammatical relation.

51.4 Differential Object Marking

51.4.1 The accusative suffix *-n*

As already mentioned, the accusative case-marking suffix *-n* (*-in* after consonant-final forms) may or may not occur on the object. The traditional observation is that only 'definite' objects can and must be marked by the accusative suffix *-n*. We use the term 'definite' here as a cover term to include a number of different types of 'definite' NPs (cf. Abbott 2004): proper names, personal pronouns, demonstrative pronouns, possessive NPs, NPs with a universal quantifier. An example of each of these categories is presented below in (15):

(15) a. *aster almaz-*(in) ayy-əčč*
Aster Almaz–ACC see.PRF-3F
'Aster saw Almaz' (Proper name)

b. *aster issu-*(n) ayy-əčč*
Aster he–ACC see.PRF-3F
'Aster saw him' (Personal pronoun)

c. *aster yih-*(in) ayy-əčč*
Aster this–ACC see.PRF-3F
'Aster saw this' (Demonstrative pronoun)

978–0–19–920647–6 52-Spencer-ch51 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 747 of 755 May 16, 2008 19:0

d. *aster liǐj-u-*(n) ayy-əčč*
Aster child-DEF–ACC see.PRF-3F
'Aster saw the child' (NPs with the definite suffix *-u*/*-wa*)

e. *aster liǐj-e-*(n) ayy-əčč*
Aster child-POSS.1SG – ACC see.PRF-3F
'Aster saw my child' (Possessed NPs)

f. *aster yə-ine-*(n) liǐj ayy-əčč*
Aster POSS-1SG-ACC child see.PRF-3F
'Aster saw my child' (Possessor NPs)

g. *aster hullu–*(n) liǐj-očč ayy-əčč*
Aster all child-PL see.PRF-3F
'Aster saw all the children' (Quantified NPs ('all'))

h. *aster iyyandandu–*(n) liǐj ayy-əčč*
Aster each child see.PRF-3F
'Aster saw each child' (Quantified NPs ('each'))

As the examples above show, in every instance the object NP must be marked by the accusative case suffix *-n* lest the construction becomes ungrammatical.

In addition to the object noun phrases exemplified above, generic noun phrases behave in a similar manner. Thus consider (16):

(16) *igziabiher səw–*(in) bə-məlk-u fət'ər-ə*
God man-ACC with-image-POSS.3M create.PRF-3M
'God created man in His image' (Genesis, 1: 27)

Notice that the direct object *səw* 'man' is generic in the context of (16) and is marked by the accusative suffix.

51.4.2 The distribution of the object/indirect object agreement suffix

The accusative suffix *-n* can also be used to mark the recipient of a three-place predicate such as 'give'. Thus compare the following constructions:

(17) a. *ləmma lə-liǐj-u məs'haf sət't'-ə-(w)*
Lemma to–child–DEF book give.PRF-3M-3MO
'Lemma gave a book to the child'

b. *ləmma lə-liǐj-u məs'haf-u-n sət't'-ə-(w)*
Lemma to–child–DEF book-DEF-ACC give.PRF-3M-3MO
'Lemma gave the book to the child'

(18) *ləmma liǐj-u-n məs'haf sət't'-ə-*(w)*
Lemma child–DEF-ACC book give.PRF-3M-3MO
'Lemma gave the child a book'

978–0–19–920647–6 52-Spencer-ch51 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 748 of 755 May 16, 2008 19:0

Notice that in (17) the recipient is marked by the preposition *lə* 'to', whereas in (18) the same argument is marked by the suffix *-n* in a construction similar to dative-shift in English.[1] What is important for present purposes is the fact that when the recipient argument does not occur with the preposition *lə-* and is instead marked by the accusative *-n*, the object/indirect object pronominal suffix (*-w*) on the verb becomes obligatory – agreeing with the recipient argument. It is thus perhaps more appropriate in such contexts to refer to this marker as a 'pronominal suffix' rather than a mere agreement marker (as also argued in Mullen 1986).

It is also worthwhile to mention here that when the recipient argument is omitted or occurs as a pronoun (particularly with first and second person pronouns) it triggers agreement even though in the latter case the recipient may occur with the preposition *lə-*, as shown in (19a):

(19) a. *ləmma lə-ɨne məs'haf sət'ət'-ə-*(ññ)*
Lemma to-me book give.PRF-3M-1SO
'Lemma gave a book to me'
b. *ləmma məs'haf sət'ət'-ə-*(ññ)*
Lemma book-DEF-ACC give.PRF-3M-3MO
'Lemma gave a book to me'

We have seen that while the object/indirect object agreement marker is often optional, there are numerous constructions where it is obligatory. Thus, whenever the direct object is ellipsed, the object pronominal suffix is obligatory. Consider (20) below, where the object agreement suffix is optional as expected (see Haile 1970 for further discussion):

(20) *almaz məs'haf–u-n wəssəd-əčč-(ɨw)*
Almaz book-DEF.M-ACC take.PRF-3F-3MO
'Almaz took the book'

If the direct object ('the book') is omitted – as it is when it is recoverable from the immediate discourse context – the object agreement marker is obligatory as shown in (21):

(21) Q. *məs'haf–u yət nəw?*
book-DEF where be.3M
'Where is the book?'
A. *almaz wəssəd-əčč-*(ɨw)*
A. take.PF-3F-3MO
'Almaz took it'

In response to the question 'where is the book?' the direct object ('the book') is omitted but agrees with the object pronominal suffix (see Haile 1970: 105). This

[1] The formal property of ditransitive verbs is much more complex. Unlike the verb *sət'ət'ə* 'to give', the closely related verb *lakə* 'to send' does not readily allow the preposition *lə-* to be substituted by the accusative suffix *-(ɨ)n*.

suggests that, to the extent that object agreement normally occurs with definite objects, ellipsed NPs are treated as definite (see also Abbott 2004).

Another construction in which object/indirect object agreement is normally obligatory is the causative. The embedded clauses in (22a) and (22b) below are transitive and intransitive respectively. We see that in both cases the causee can take the accusative marker *-n* and the verb occurs with an object agreement marker cross-referencing the causee.

(22) a. *aster ləmma-n məs'ihaf ind-i-gəza*
Aster Lemma-ACC book COMP-3M-buy.IPFV
a-dərrəg-əčč-iw
CAUS-make.PRF-3F-3MO
'Aster made Lemma buy a book'

b. *aster ləmma-n wədə bet ind-i-hed*
Aster Lemma-ACC to home COMP-3M-go.IPFV
as-gəddəd-əčč-iw
CAUS-force.PRF-3F-3MO
'Aster forced Lemma to go home'

To summarize, when a direct object is definite it is marked by the accusative case *-n*, and can optionally trigger agreement on the verb. The set of forms that are obligatorily marked by the accusative case consists of definite NPs, including demonstrative pronouns, personal pronouns, proper names, as well as quantified and generic NPs.

In the next section we investigate the roles played by the overt accusative case and object/indirect object morphology on the verb in the phenomenon of differential subject marking.

51.5 Differential Subject Marking

In this section I show that the quirky (non-canonical) case-marking of certain subjects by and large occurs with a semantically homogeneous group of verbs – broadly construed as EXPERIENCER predicates.

As already mentioned, there is enough evidence to establish the existence of the subject grammatical relation in Amharic. Typically, subjects occur in clause-initial position, get the nominative case, and obligatorily agree with a pronominal marker on the verb. However, there are some constructions that are problematic in that the 'subject' appears to lack some or all of these coding properties. Consider the following examples:

978–0–19–920647–6 52-Spencer-ch51 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 750 of 755 May 16, 2008 19:0

(23) *aster tə-č'ənnək'-əčč*
Aster INCH-worry.PRF-3F
'Aster is worried'

(24) *aster tə-dənnək'-əčč*
Aster INCH-astonish.PRF-3F
'Aster is astonished'

In both (23) and (24) the subject (S) of the intransitive clause occurs in clause-initial position and agrees with the pronominal suffix on the verb as expected. Now consider the synonymous constructions in (25) and (26) below:

(25) *aster-(in) č'ənnək'-at*
Aster-(ACC) worry.PRF.3M-3FO
'Aster is worried'

(26) *aster-(in) dənnək'-at*
Aster-(ACC) astonish.PRF.3M-3FO
'Aster is astonished'

While (25) and (26) have the same meaning as (23) and (24) respectively, there are important formal differences between the two sets of constructions. First, in (25) and (26) the single argument of the verb – the 'subject' – is optionally marked by the suffix *-n* which is normally found with the direct object grammatical relation. Second, the single argument of the verb – semantically an Experiencer – obligatorily agrees with a pronominal suffix on the verb that normally occurs with objects. The subject pronominal suffix on the verb has the features third person, masculine, and singular. Since subject agreement on the verb is obligatory in Amharic, the question becomes: what argument does the subject agree with in (25) and (26)?

Notice that the presence of the accusative case marking on the subject noun phrase in (25) and (26) is unusual given that normally only (definite) objects of a transitive verb take the *-n* suffix which marks the accusative case. This is shown again in (27), where object noun phrases with the patient role are marked by the accusative case:

(27) *səw-ɨyyəw aster–*(ɨn) sam-(at)*
man-DEF Aster-(ACC) kiss.PRF.3M-3FO
'The man kissed Aster'

While the accusative suffix *-ɨn* is obligatory in (27), the object pronominal suffix is not. The reader will have noticed that this scenario is the converse of what we find with the constructions in (25) and (26) where the accusative case marker on the Experiencer is optional but the object pronominal suffix on the verb is obligatory.

Constructions in which the single argument of an intransitive verb agrees with an object pronominal suffix as in (25) and (26) are not uncommon in Amharic. However, before discussing these constructions, a more detailed description of Amharic psychological (experiencer) predicates is in order.

51.5.1 Experiencer Predicates

Following previous work (cf. Amberber 1996, 2000, 2001, 2002), I assume that there are two major types of experiencer predicates in Amharic: Subject Experiencer (SubjExp) predicates and Object Experiencer (ObjExp) predicates (the terms SubjExp and ObjExp are adapted from Pesetsky 1995 and related work). These are illustrated in (28)–(29):

(28) *aster tə-dəssət-əčč*
Aster INCH-happy.PRF-3FS
'Aster is happy'

(29) *sit'ota-w aster-in as–dəssət-at*
gift-POSS.3M Aster-ACC CAUS-INCH.happy.PRF.3M-3FO
'His gift made Aster happy'

In (28a) and (29) the Experiencer argument is in subject position (agrees with the subject pronominal suffix on the verb). The key difference between (28) and (29) is transitivity: the former is intransitive whereas the latter is transitive. Au: Please check (28a)

In (29), on the other hand, the Experiencer is in object position (is marked by the accusative suffix *-n*, and agrees with the object pronominal suffix on the verb). Note that in (29) the verb is causativized by the prefix *as-* (as opposed to the prefix *a-*) that is normally used to causativize transitive verbs or unergative intransitives.[2] The construction in (29) becomes ungrammatical with the causative prefix *a-*:

(29′) **sit'ota-w aster-in* ***a–****dəssət-at*
gift-poss.3m Aster-ACC **CAUS**-BE.happy.PRF.3M-3FO

Thus, it appears that the causativization pattern of the intransitive experiencer predicate is the same as the causativization pattern of unergative intransitives as shown in (30) below with the unergative intransitive verb *č'əffərə* 'dance':

(30) a. *aster č'əffər-əčč*
Aster dance.PRF–3FS
'Aster danced'
b. *ləmma aster-in* ***as–****č'əffər-at*
Lemma A.-ACC **CAUS**-dance.PRF.3M-3FO
'Lemma made/let Aster dance'
c. **ləmma aster-in* ***a–****č'əffər-at*
Lemma A.-ACC **CAUS**-dance.PRF.3M-3FO

This constrasts with the situation of unaccusative intransitives where the verbs take the prefix *a-* in causativization. Thus, consider the following example:

[2] We note here that while this is the normal pattern, i.e. the prefix *a-* occurs mostly with (unaccusative) intransitives, there are some exceptions. In some cases the exception is due to a phonological constraint; when the basic stem starts with the vowel *a-* it does not take the prefix *a-* as its causative marker but rather the prefix *as-*.

978–0–19–920647–6 52-Spencer-ch51 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 752 of 755 May 16, 2008 19:0

(31) a. *aster məťťa-čč*
Aster come.PRF–3FS
'Aster came/arrived'
b. *ləmma aster-in a–məťťa-at*
Lemma Aster-ACC CAUS-come.PRF.3M-3FO
'Lemma brought Aster'

This raises the interesting question: why do intransitive experiencer predicates behave like unergative verbs in their causativization pattern? Is it an accident that intransitive experiencer predicates, unergatives, and transitive predicates do not all take the causative prefix *a-* in causativization?

I argue that these verbs behave similarly with respect to the morphological causative because all are instances of a two-place predicate. Thus, both types of intransitive predicate (unergatives and intransitive experiencer predicates) have underlyingly two arguments. Following the suggestion made by Hale and Keyser (1993, 2002), I assume that unergative verbs appear to have a single argument because the internal argument – sometimes referred to as 'cognate' object – is not normally overtly expressed. In the case of intransitive subject experiencer predicates of the type like 'be happy', 'be annoyed', etc. one can assume that there is an underlying causer argument that can be referred to by the term AMBIENT-CAUSER (A-causer) as in Pesetsky (1995: 112). It may be this argument that controls subject agreement in constructions such as (25)–(26). When this argument is 'suppressed' (by a morphological process similar to reflexivization/passivization) we get constructions such as (23)–(24) where the Experiencer occurs in the nominative.

The fact that subject experiencer predicates exhibit quirky Case properties cross-linguistically is by now well documented (cf. Andrews 2001, Aikhenvald et al. 2001, Bhaskararao and Subbarao 2004, among others). According to McCawley (1976) indirect subject constructions involve verbs that express events such as those listed in (32):

(32) a. sensory and mental experience
b. emotional experience
c. physical and biological experience
d. need/duty/obligation
e. possession/existence
f. happenstance

It is observed that intransitive subjects which are more 'affected', or which have less 'control' over the event, in the sense of Hopper and Thompson (1980), are likely to be marked by the accusative/dative, whereas intransitive subjects which are 'agentive', or exert more control over the event, may be marked by the nominative. For a more elaborate semantic classification of verbs which exhibit differential

marking ('non-canonically marked arguments') see Onishi (2001: 23ff); see also Malchukov (2005) for relevant discussion.

In Amharic, there are at least two broad classes of verbs that exhibit this kind of differential behaviour. The first class involves verbs that can be characterized as experiencer predicates – broadly construed involving not only PSYCHOLOGICAL experience (as in (25)–(26)) but also SENSATION and other (physical and mental) experiences as in (33)–(35):

(33) *aster-(in) t'əmm-at*
Aster-(ACC) thirst.PRF.3M-3FO
'Aster is thirsty'

(34) *səww-očč-u-(n) rab-ačč-əw*
man-PL-DEF-(ACC) hunger.PRF.3M-3PLO
'The men are hungry'

(35) *aster-(in) ammə̀m-at*
Aster-(ACC) sick.PRF.3M-3FO
'Aster is sick'

Notice that in (33)–(35) there is obligatory object agreement with the argument that undergoes some sort of sensation, just as with the argument that undergoes psychological experience in (25)–(26). Subject agreement is with a third person masculine argument.

Now if we assume that the non-canonically marked experiencers in (33)–(35) are similar to the non-canonically marked experiencers of the psychological predicates in (25)–(26), we would predict that these subjects will exhibit canonical behaviour (i.e. will be marked as nominative and control subject agreement) when the verb is inflected with the prefix *t(ə)-* similar to (23)–(24). This is exactly what we find, as shown in (36)–(38):

(36) *aster tə-t'əmma-čč*
Aster INCH-thirsty.PRF-3F
'Aster is thirsty'

(37) *səww-očč-u tə-rab-u*
men-PL-DEF INCH-hungry.PRF-3PL
'The men are hungry'

(38) *aster t-ammə̀m-əčč*
Aster INCH-sick.PRF-3F
'Aster is sick'

The class of temperature verbs such as *bərrədə* 'it has become cold', behave in a similar manner. Although these verbs are typically used to express temperature, as shown in (39), they can also be used with an experiencer argument, as shown in (40):

(39) *yi-bərd-all*
3M-IPFV.be.cold-AUX.3M
'It is cold'

(40) *aster-(in) bərrəd-at*
Aster-(ACC) be.cold.PRF.3M-3FO
'Aster is cold'

As in the psychological predicates, there is an obligatory object agreement with the argument that is experiencing the temperature.

Note that semantically the Experiencer argument of a typical experiencer predicate exerts little or no control over the event. For example, with the verb in (41) below the event is conceptualized as something that takes place without volition, and thus the Experiencer argument is agreed with by an object (indirect object) pronominal suffix:

(41) *rabə-ññ*
hunger.PRF.3M-1SO
'I am hungry'
(lit. 'it hungers me')

To summarize, constructions that exhibit differential subject properties are, to a large degree, semantically homogeneous. The argument which is optionally marked by the accusative case suffix *-n* and agreeing with an apparent object pronominal suffix is often some kind of experiencer in different domains – including those undergoing emotional and physical sensations.[3]

51.6 Conclusion

The Amharic verb is obligatorily marked for subject agreement. It is optionally marked for object and indirect object agreement. The subject is non-overtly marked for nominative case, and the object is marked for accusative case (by the suffix *-n*) if it is definite.

[3] One way to provide a unified account for such constructions is to assume that the predicates in question allow the Experiencer argument to be optionally topicalized. In such topicalized contexts the agreement marker on the verb functions as an anaphoric pronominal suffix. This is in line with the observation that in all constructions which involve movement – such as relativization, *wh*-questions, and clefts – the object pronominal suffix is obligatory on the verb and functions as an anaphoric pronoun rather than a mere agreement marker. Due to space constraints these issues will not be explored in this chapter (but see Amberber 2005 for relevant discussion).

978–0–19–920647–6 52-Spencer-ch51 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 755 of 755 May 16, 2008 19:0

Amharic, like many other languages, exhibits the phenomenon of non-canonical marking of the subject grammatical relation. First, while normally the subject is assigned nominative case and controls subject agreement, in some constructions it exhibits an optional accusative marking and agrees with an obligatory object/indirect object pronominal suffix on the verb. Predicates that exhibit such properties include verbs of emotion, physical sensation, and temperature, among others. We have argued that intransitive experiencer predicates such as 'worry' are conceptually more complex, with an abstract causer argument (Ambient Causer) which agrees with the subject pronominal suffix, whereas the experiencer argument agrees with the object/indirect object pronominal suffix on the verb. The fact that the experiencer argument obligatorily agrees with the object/indirect object pronominal suffix on the verb was accounted for by motivating a movement analysis where the experiencer argument is left-dislocated into an adjunct position. The pronominal suffix on the verb is then analysed as an anaphoric pronoun rather than a mere agreement marker. Support for this analysis comes from clefts, relative clauses, and *wh*-constructions, which all require the presence of the object/indirect object pronominal suffix on the verb.

It is hoped that this chapter will contribute to the growing body of empirical work on non-canonical subject marking. The analysis presented in the chapter shows how non-canonical behaviour is an artefact of an alternative way of encoding grammatical relations which is predictable from a semantically homogeneous class of verbs.

CHAPTER 52

CASE IN AN AUSTRALIAN LANGUAGE

DISTRIBUTION OF CASE AND MULTIPLE CASE-MARKING IN NYAMAL

ALAN DENCH

Nyamal is an Australian language of the Pama-Nyungan family, originally spoken in the Pilbara region of Western Australia, around the present town of Marble Bar.[1] Like most Pama-Nyungan languages, Nyamal is a highly agglutinating dependent-marking language, with a wealth of suffixal nominal inflection. In Nyamal, as in many Australian languages, nominal suffixes serve a wide range of dependency-marking functions, which include but are not restricted to traditional 'case'.

[1] Fieldwork on Nyamal was conducted in 1993 and 1995 and the author is especially grateful to †Daisy Williams for teaching him what he knows of the Nyamal language. Only a handful of speakers of Nyamal remain. The research reported here was supported by the Australian Research Council (A59131653, A59532829). My thanks to the editors, Edith Moravscik, and an anonymous referee for helpful suggestions. The usual disclaimers apply.

Nyamal has a very complex system of case-marking. This complexity is due to its very strong tendency towards multiple case-marking combined with variation in the case-marking selected by arguments of different predicates, in different clause types, and by different patterns of case syncretism in different classes of nominals. This chapter describes and exemplifies that complexity. Section 52.1 sets out the variables affecting the choice of case-marking for arguments in clauses, 52.2 describes the patterns of multiple case-marking.

52.1 Variables determing clause level case marking

Nyamal is described as a language with split-ergative case-marking and thus its patterns of marking can be usefully considered in relation to typological frameworks advanced for the description of split-ergative alignment systems. For example, Dixon (1994: 70) argues that different split systems are conditioned by one or more of four parameters; the semantic nature of the main verb, the semantics of the core NPs, tense/aspect/mood categories within the clause, and the dependency status of the clause. Dixon's approach is influenced by Silverstein's (1976, 1993) outline of a general model of case-marking. Silverstein considers case-marking in any particular language to result from the interaction of four independent variables: case relations (or predicate-argument relations), the inherent referential content of nominals (noun phrases and/or lexical items), clause-linkage type, and discourse coreference relations (e.g. switch reference).[2] Essentially, all of these factors are relevant to description of case assignment in Nyamal. These are classed into three broad parameters here: predicate type, nominal type, and clause type – where clause type is a function of verbal inflection (tense/aspect/mood, switch reference) and dependency status (see Table 52.3 below).

To begin with, case marking in a clause depends on the **class of predicate**. Six broad classes of argument-taking predicate can be recognized for Nyamal. These types, together with their associated case frames and reference to illustrative examples[3] (which follow) are shown in Table 52.1.[4]

[2] Silverstein's third variable can be generalized to allow the categorization of 'independent' clauses with particular temporal values or modalities (thus naturally incorporating Dixon's tense and aspect based splits) if clause-linkage type is understood in terms of the logical relations which hold between a clause and the discourse more generally.

[3] For abbreviations used in examples see the list at the front of the book.

[4] Constituent order in Nyamal is relatively free. There is a tendency towards a basic VO order, with subjects either in an initial topic position or following the verb (and usually its complement).

978–0–19–920647–6 53-Spencer-ch52 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 758 of 769 May 16, 2008 19:24

Table 52.1. Nyamal predicate types

Predicate type	Case frame	Example	
intransitive nominal	NOM	muri 'lame'	(1)
intransitive verb	NOM	nyina-Ø 'sit, stay'	(2)
extended intrans. nominal	NOM DAT	miranu 'know'	(3)
extended intransitive verb	NOM DAT	wurtama-L 'wait for'	(4)
transitive verb	ERG ACC	nyanyja-L 'bite'	(5)
ditransitive verb	ERG ACC DAT	manya-L 'give'	(6)
	ERG ACC LOC	juntiya-L 'tell'	(7)

(1) *Nyaa yukurru muri.*
this.NOM dog.NOM lame
'This dog is lame.'

(2) *Jilya-kujalpa nyini-yampa-pulu.*
child-DU.NOM sit-PRES-3DU
'The two children are sitting down.'

(3) *Ngaja para miranu nyaa-yu jarrunpa-ku.*
1SG.NOM 3SG.DAT know this-DAT man-DAT
'I know this man.'

(4) *Ngaja wurtama-lka-rna para nyaa-yu*
1SG.NOM wait for-PRES-1SG 3SG.DAT this-DAT
'I'm waiting for this one.'

(5) *Maruntu nyanyja-la-ya yukurru-lu.*
goanna.ACC bite-PRES-3PL dog-ERG
'Dogs catch goannas.'

(6) *Mantu manya-rna para ngajuku-lu mirtari-lu kamparra-ku*
meat.ACC give-PAST 3SG.DAT 1SG.GEN-ERG aunty-ERG little.one-DAT
jilya-yu.
child-DAT
'My aunty gave meat to the little child.'

(7) *Mirtanya-lu juntiya-rna pulara kamparra-la wangka.*
old man-ERG tell-PAST 3SG.LOC little.one-LOC story.ACC
'The old man told the little child a story.'

As (6) and (7) illustrate, ditransitive verbs typically select either the dative or locative as the case of the recipient or beneficiary. There is some variation for particular verbs. For example, while *juntiya*-L, 'tell', always selects the locative, *manya*-L, 'give' may select either the dative or the locative.

Table 52.2. Case syncretisms in Nyamal

nominals	pronouns	3sg pronoun
ERGATIVE	ERGATIVE	ERGATIVE
NOMINATIVE	NOMINATIVE	NOMINATIVE
ACCUSATIVE	ACCUSATIVE	ACCUSATIVE
GENITIVE	GENITIVE	GENITIVE
DATIVE	DATIVE	DATIVE
ALLATIVE	ALLATIVE	ALLATIVE
LOCATIVE	LOCATIVE	LOCATIVE
ABLATIVE	ABLATIVE	ABLATIVE

In addition to there being predicate types which subcategorize for a dative argument, dative noun phrases denoting beneficiaries, or more generally 'ethical datives', can be quite freely added to a range of clause types. The syntactic status of these additional arguments is more immediately obvious in languages in which they are cross-referenced in the agreement system, such as Nyangumarta (the northeastern neighbour of Nyamal) or Warlpiri.[5] Nyamal has subject agreement in finite verbal clauses[6] but lacks agreement for non-subject arguments and so there is no straightforward test for the complement or adjunct status of these arguments. However, since there is a strong tendency for such arguments to be introduced by the use of a dative (or locative) pronoun in postverbal position, there is a prima facie case for considering them to have much the same status as subcategorized arguments. True adjuncts do not typically appear with the appropriately case-marked pronouns.

As in most Australian languages described as split-ergative, the choice of case inflection depends partly on the **class of nominal**. In Nyamal, pronouns inflect on a nominative–accusative pattern while other nominals inflect on an ergative–absolutive pattern. Following Goddard (1982), this split is treated as a syncretism of forms across three cases – ergative (the case of transitive subject), nominative (the case of intransitive subject), and accusative (the case of transitive object). Thus nominals generally show a syncretism of nominative and accusative while pronouns show a syncretism of ergative and nominative. The two classes also show a range of additional syncretisms as summarized in Table 52.2 – cases in the same box are syncretized. The third person singular pronoun alone codes a distinction between genitive and dative.

[5] Simpson (1991) presents a detailed description, and formal account, of Warlpiri dative 'External Objects'. Additional complement datives in Nyangumarta are described in Sharp (2004).

[6] Subject agreement is clear in examples (2), (4), and (5). Third person singular subject agreement is not marked, as can be seen in examples (6) and (7).

978–0–19–920647–6 53-Spencer-ch52 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 760 of 769 May 16, 2008 19:24

The syncretisms in Table 52.2 are easily established. In Nyamal, as in most Australian languages, part–whole possessive relationships involve no overt indication of possession; nominals denoting the part and the whole are simply juxtaposed. Thus the syncretisms can be shown by comparing noun phrases which include a pronoun denoting the whole and a nominal denoting the part; both pronoun and nominal have the same case function (Goddard 1982).

In his discussion of the case assignment patterns of dependent clauses in split-ergative languages, Silverstein (1976) describes patterns in which the 'plain' ergative–absolutive marking pattern of transitive predicates is suspended. His Variable III, **clause-linkage type**, involves an implicational hierarchy of logical relations between clauses sharing a coreferential noun phrase. This hierarchy predicts the relative likelihood of particular case-marking patterns appearing in particular kinds of clauses. For (split) ergative languages, in moving up the hierarchy 'it becomes more and more the case that a language will suspend the lexical hierarchy for split ergative, use anti-passivised forms of transitives in nominative–dative "normal" forms, and nominalise with a possessive or equivalent schema' (Silverstein 1976: 163).

Case-marking in Nyamal non-finite clauses conforms to this prediction. For example, the privative suffix, *-yapa*, is typically associated with nominals but is also used as an inflection on verbs (with conjugation-dependent forms, *-yapa* and *-lapa*) to denote a negative state. Transitive verb stems marked with the privative inflection select a nominative–dative case frame, as the following examples show.

(8) *Paja-lapa ngaja mantu-yu, yirra-yapa ngaja.*
eat-PRIV 1SG.NOM meat-DAT tooth-PRIV 1SG.NOM
I don't eat meat, I have no teeth.

(9) *Nyaa yanga-lapa mama-yu. Palura wilyuru-ngarri-yampa.*
this.NOM follow-PRIV father-DAT 3SG.NOM not.want-INCH-PRES
'This one doesn't follow his father. He doesn't like to.'

Transitive verbs marked with the purposive inflection may appear in non-finite clauses either as independent desideratives or as negative commands, in which cases they select an ergative–accusative case frame. But they also occur in dependent implicated purpose clauses in which case their apparent objects are marked with the dative. First, (10) illustrates both a desiderative and negative command. Here, the ergative secondary predicate[7] on the transitive subject in the first clause and the accusative pronoun in the second clause make the ergative–accusative frame quite clear. By contrast, (11) involves a purpose clause with a dative object.

[7] For a recent discussion of secondary predicates of this kind see Schultze-Berndt and Himmelmann (2004).

(10) *Ngatha kama-larta mungu-ngku. Mirta nganya wangkayinya-larta*
1SG.ERG cook-PURP alone-ERG not 1SG.ACC scold-PURP
nyunta. Ngatha kama-larta mungu-ngku.
2SG.ERG 1SG.ERG cook-PURP alone-ERG
'I can cook it on my own. Don't you tell me (what to do). I can do it on my own.'

(11) *Ngunti-rna-rna jilya kurti-larta yurta-yu.*
send-PAST-1SG child.ACC get-PURP.ACC fish-DAT
'I sent the child to get fish.'

Nominalized clauses in Nyamal show more complicated patterns of case modification, described as 'associating case' (following Evans 1985,1995a; Dench and Evans 1988). In the simplest patterns, the nominalizing suffix that appears on the verb also appears on arguments of that verb. The usual case-marking patterns are suspended. Example (12) illustrates this for the intentive nominalization. The nominalized verb appears with what would ordinarily be a dative complement, in the first clause, and what would be the ergative subject, in the second clause.

(12) *Nyaa mantu yapa-lwanti jilya-karrangu-wanti, paja-lwanti*
this.NOM meat.NOM give-INTENT child-PL-INTENT eat-INTENT
jilya-karrangu-wanti.
child-PL-INTENT
'This meat is (intended) for giving to the children, for the children to eat.'

Perfective relative clauses involve a verb inflection which is cognate with the ablative nominal suffix. In simple transitive subordinate clauses, the ablative suffix appears on the (otherwise unmarked) object of the nominalized verb as an associating case (13). In (14) the ablative follows a locative-marked adjunct in the nominalized clause. Here the associating case suffix follows rather than replaces the usual argument case-marking. Note also that the subordinate clause in (14) is (un)marked nominative in agreement with its head. In (13) the subordinate clause is marked dative in agreement with its head. This 'complementizing' case-marking is discussed in section 52.2 below.

(13) *Wajarri-nya-rna para jarrunpa-ku, kama-lara-ku*
look.for-PAST-1SG 3SG.DAT man-DAT cook-PERFREL-DAT
mantu-kulyara-ku.
meat-ABL-DAT
'I looked for the man, after he cooked the meat.'

(14) *Punpal-ngarri-nya-rna kuji kayarri-kulyara piju-ngka-kulyara.*
sore-INCH-PAST-1SG leg.NOM swim-PERFREL.NOM river-LOC-ABL.NOM
'My leg is sore from swimming in the river.'

While the patterns illustrated in (12), (13), and (14) might be easily interpreted as the simple spreading of an inflection from the head of a nominalized clause

to its dependents, just as case is distributed within noun phrases, other clause types are more complicated in their selection of associating cases. In particular, the same-subject and different-subject relative clauses select associating case suffixes that are not related to the verbal inflection. The usual associating case found in different-subject relative clauses is *-kapu*, SOURCE. In (15) this suffix is attached to the bare stem of the object but follows the usual case-marking of the locative complement. Same-subject relative clauses select the ACTIVITY suffix as an associating case.[8]

(15) *Ngaja wiya-lka-rna ngunya jarrunpa, juntiya-nyjanu*
1SG.NOM see-PRES-PAST that.ACC man.ACC tell-RELDS.ACC
ngajala-kapu wangka-kapu.
1SG.LOC-SRC.ACC word-SRC.ACC
'I can see that man who told me the story.'

In summary there are three different patterns, which can be identified best by their consequences in marking the arguments of transitive verbs: ergative–accusative marking is found in finite declarative clauses, nominative–dative is found in some non-finite clause types, and a variety of associating case-marking patterns occur in nominalized clauses. Table 52.3 lists different clause types – identifiable by verbal inflectional category – and the range of case patterns found in each. Some clause types show more than one pattern of marking; the less common pattern is indicated in parentheses.

Table 52.3. Nyamal case-marking patterns by clause type

Clause type	Verb inflection	Case-marking		
		ERG-ACC	NOM-DAT	ASSOCIATING
Declarative	*various*	✓		
Desiderative	purposive	✓	(✓)	
Negative generic	privative		✓	
Implicated purpose	purposive		✓	
Imperfective relative	relative (SS, DS)		(✓)	✓
Jussive complement	purposive		(✓)	✓
Perfective relative	perfective			✓
Nominalization	*various*			✓

[8] The *-karra* ACTIVITY suffix is attached to nominals that are seen as characterizing a particular kind of activity taking place in a time interval overlapping that of the main clause predication. In almost all cases of its use in finite main clauses the suffix occurs in subject-oriented second predications.

52.2 Multiple case-marking

Nyamal exhibits a high degree of multiple case-marking. Nominal suffixes serve a range of functions beyond the traditional marking of the arguments of predicates at the level of the clause, coding predicate argument relations within a variety of constituent types. Thus a particular word may bear a set of case suffixes, each indicating its role in successively higher constituent structures. The complexity of such multiple case-marking systems in Australia is explored in Dench and Evans (1988). A wider cross-linguistic survey of such multiple case-marking is to be found in the set of papers collected together in Plank (1995a) and from this it is clear that in most parts of the world multiple case-marking is restricted to just two levels – hence double case – and prototypically involves the further inflection of a genitive-marked possessive modifier within a noun phrase (Suffixaufnahme).

The more extreme examples of multiple case-marking thus appear to be restricted to just two groups of languages: the Australian languages of the Pilbara region of Western Australia and of the Tangkic group spoken in north Queensland (e.g. Kayardild, described by Evans 1995a). Nyamal is the most elaborated of the Pilbara languages. The following subsections present the different parameters by which the patterns of multiple case-marking in Nyamal can be described.

52.2.1 Functions of case suffixes

Nominal suffixes in Nyamal are found with all five of the different functions identified in Dench and Evans (1988). These are:

a. *relational*: the prototypical function of case-marking to code argument roles at the clause level,
b. *adnominal*: coding relationships between nominals within the one noun phrase,
c. *referential*: nominal adjuncts of various kinds, including secondary predicates, are formally linked to the arguments of the main predicate in a clause through case agreement,
d. *associating*: linking arguments to nominalized verbs in subordinate clauses, and
e. *complementizing*: the use of case morphemes to mark a dependency relationship between clauses – either C-complementizing, linking an argument of the subordinate clause to a coreferential NP in the main clause, or T-complementizing, indicating a temporal or logical connection between the two clauses.

Different nominal suffixes have different ranges of use. The range of functions of (a selection of) Nyamal nominal suffixes is presented in Table 52.4.

The particular functions of the ergative are predictable given its restricted role as the marker of transitive subjects – it also marks adjuncts of transitive subjects and subordinate clauses controlled by a matrix transitive subject. The dative is

Table 52.4. Range of functions of Nyamal nominal suffixes

	Relational	Adnominal	Referential	C-comp	T-comp	Associating
Ergative	✓		✓	✓		
Dative	✓	✓	✓	✓		
Locative	✓	✓		✓	✓	
Allative	✓	✓				
Ablative	✓	✓				✓
Intentive	✓	✓				✓
Source		✓				✓
Activity		✓				✓
Privative		✓				
Proprietive		✓				

similar except that the one suffix form codes both dative and genitive case, hence the adnominal uses of the suffix as a genitive. The locational cases – locative, allative, and ablative – mark both locational adjuncts and adnominal relations within noun phrases. The locative also marks subordinate clauses controlled by a matrix locative adjunct, marks temporal adverbial clauses which describe some event taking place at the time of the matrix clause event, and serves as the default complementizing case suffix on subordinate clauses that do not share an argument with their controlling clause. The ablative and allative do not have these additional functions. The nominal suffixes used as associating cases (intentive, source, activity; see Table 52.4) have primarily adnominal functions. Other adnominal suffixes, such as the proprietive and privative, have only an adnominal function.

Example (16) illustrates a number of the functions identified in Table 52.4. To begin with, *kurtanpa* 'bag' (subscript 1), is marked with a relational locative suffix. In the second line, *puja-ngka* 'on back' (2), involves an adnominal locative and the expression functions as a second predication on the (unmarked accusative) object of the verb 'carry', that is, the bag of pearlshell. Similarly, the adverbial nominal, *karlinypa* 'returning' (3), is a second predication on the object of 'carry'. By contrast, the second predications in the third line (4, 5) are linked to the transitive subject of 'carry' by referential ergative case. The first of these, 'from the sea' (4), involves adnominal locative and then ablative suffixes preceding the referential ergative suffix. The adverbial nominal, *karlinypa*, 'returning' (5), simply bears the referential ergative suffix. Finally, the third line includes a simple relational allative expression (6). As a clause-level locative adjunct, this allative expression does not agree with the ergative subject.

(16) *i. Pirrapirra winyaya-lamu-rtu kurtanpa-la$_1$, pirrapirra.*
pearlshell.ACC fill-USIT-1PL.EXCL bag-LOC pearlshell.ACC

978–0–19–920647–6 53-Spencer-ch52 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 765 of 769 May 16, 2008 19:24

ii. *Puja-ngka$_2$ katinya-lamu-rtu-nga puja-ngka$_2$ karlinypa$_3$,*
back-LOC.ACC carry-USIT-1PL.EXCL-NOW back-LOC.ACC returning.ACC
iii. *kujungurru-la-kulyara-lu$_4$ now, karlinypa-lu$_5$, yurlu-karni$_6$.*
sea-LOC-ABL-ERG now returning-ERG camp-ALL
'We would fill pearlshell into the bag. Now we would bring it back, on our backs, as we came back from the sea, now, to camp.'

Examples (17) and (18) illustrate the complementizing uses of nominal suffixes. In (17), the locative functions as a T-complementizing case on the two temporal clauses. In the first clause, the locative complementizing case suffix occurs on both the verb, 'fall', and its subject, 'sun'. In (17), the subordinate purpose clause is controlled by a locative noun phrase in the first clause and thus the subordinate clause verb bears a locative C-complementizing case.

(17) *Yarnta-ka warni-nyjanu-la, warrukarti-ngarri-nyjanu-la,*
sun.NOM-LOC fall-RELDS-LOC dark-INCH-RELDS-LOC
karlinyjarri-yampa-rna yurlu-karni.
return-PRES-1SG camp-ALL
'When the sun goes down, when it starts to get dark, I go home.'

(18) *Nyuntala ngurluma-lka-rna wangka wanyarrima-larta-la.*
2SG.LOC give-PRES-1SG word.ACC hear-PURP-LOC
'I'm giving you words (for you) to listen to.'

Given the patterns of distribution of case suffixes, words can bear a number of suffixes each with a different function. The patterns permit the construction of nominal expressions bearing up to four levels of case-marking. These are attached to a stem in a sequence reflecting their structural level. Thus a maximally inflected stem may bear suffixes encoding adnominal, relational, associating, and complementizing functions.

52.2.2 Morphological coding conventions

Morphological coding conventions determine the distribution of case suffixes to constituents within a phrase or clause marked for case. The principal coding convention operating in Nyamal is 'complete concord' (Dench and Evans 1988). This pattern is clearest in noun phrases – every constituent is inflected for the particular case. An example is the ergative-marked proprietive secondary predicate in (19), which is interpreted here as coding the instrument.

(19) *Nyunanga-karta-lu pinyjil-karta-lu yaji-la-ngka-mu.*
2SG.GEN-PROPR-ERG pencil-PROPR-ERG write-ANT-2SG-ANT
'You'll write (it) down with your pencil'

Complete concord holds in relative and nominalized clause types where these function as subordinate clauses and so bear complementizing case; the complementizing case suffix appears on all constituents of the clause. In (20), the dative complementizing case suffix occurs on the subordinate clause verb and also on the object of this verb, following the SOURCE associating case suffix (and see (17) above).

(20) *Wurtama-la nyumpalanga-mu mayi-kapu-ku kama-njanu-ku.*
wait.for-ANT 2DU.DAT-ANT food-SRC-DAT cook-RELDS-DAT
'He'll wait for you two (who are) cooking food.'

The pattern of distribution for associating case appears, at first, not to fit the rule of complete concord. Examples occur in which the expected associating case suffix is not distributed to the apparent subject of the nominalized verb. In (17), for example, the subject of the first subordinate clause, 'sun', bears the locative complementizing case but does not carry the associating case suffix expected on arguments of different-subject relative clauses. This can be compared with (15) where the object of the subordinate verb bears associating case. Notice in (15) that the head (subject) of the relative clause is an accusative object of the matrix clause. I suggest that in (17), the apparent subject of the first subordinate clause be treated similarly as a (raised) locative adjunct of the matrix clause. There are examples in which the subordinate subject does bear associating case. In (21), the head of the relative clause is the object, and the subject in the subordinate clause bears associating case (and see (12) above). Example (22) shows that secondary predicates (here the adverbial nominal, *warlkanti*, 'sitting down') on the subject of a dependent clause bear the expected associating case.[9]

(21) *Kamparra-lu paja-la mantu kama-njanu ngajuku-kapu*
little.one-ERG eat-PRES meat.ACC cook-RELDS.ACC 1SG.GEN-SRC.ACC
mirtari-kapu.
aunty-SRC.ACC
'The little one is eating the meat my aunty cooked.'

(22) *Pala wiya-lka-rna warlkanti-kapu nyini-nyjanu.*
that.ACC see-PRES-1SG sitting.down-SRC.ACC sit-RELDS.ACC
'I'm watching that one, (who is) sitting down.'

52.2.3 Morphological sequence constraints

Patterns of suffix distribution predicted by the use of case suffixes at different functional levels and by the rules of concord are modified by certain morphological

[9] These patterns are similar to but ultimately differ from those in Kayardild. Evans (1995a) shows that the proper domain of associating case in Kayardild excludes the subject (and its adjuncts) in the subordinate clause.

sequence constraints. Languages in the Pilbara region show differing degrees of multiple case-marking and share very similar constraints. These are essentially of two kinds. First, most languages prevent the further inflection of nominal expressions bearing particular case suffixes — typically the ergative, accusative, and/or dative. Second, most languages place a ban on a sequence of identical suffixes.

In Nyamal, the clearest instance of a ban on the further inflection of a particular suffix is the inability of any case to follow the dative. Given that this suffix codes genitive as well as dative case, it might be expected that whenever a genitive modifier occurs in any marked noun phrase, the (adnominal) dative/genitive suffix would be followed by a relational case suffix. Such constructions occur in Thalanyji and Jiwarli, spoken in the Southern Pilbara region (Austin 1995), but do not occur in Nyamal.[10] They are avoided through the use of genitive pronouns as anaphors. That the dative cannot be further inflected is supported by the fact that dative objects in subordinate clauses bear neither an associating case inflection nor any complementizing case suffix. In (23), the dependent purpose clause is controlled by the ergative subject of the matrix clause and has ergative complementizing case on the purposive-marked verb. However, the dative-marked object of the dependent verb bears no further inflection.

(23) *Ngunya-ngku mangkurla-lu warnta kurti-la punga-lartara-lu*
that-ERG woman-ERG stick.ACC get-PRES hit-PURP-ERG
yukurru-ku.
dog-DAT
'That woman is getting a stick to hit the dog.'

Since both pronouns and nominals display a genitive/dative syncretism, it ought to follow that dative pronoun objects in purpose clauses would take further dative inflection; genitive/dative pronoun forms can bear further inflectional suffixes, as (21) shows.[11] However, there are very few examples in the corpus in which a pronoun is unambiguously the object of a purpose clause. In most examples, genitive/dative pronoun forms denoting the object of the purpose clause can be analysed as 'external dative' arguments of the matrix clause. This analysis is strongly favoured in the following two examples.

[10] Nyamal shares the constraint that the suffix, *-ku*, not be followed with a number of its southern neighbours, for example Panyjima, Yindjibarndi, and Ngarluma. In these languages, the original suffix is more widespread since it has generalized to accusative as well as dative function. However, these languages do not use the suffix to code genitive case; a new genitive suffix has been innovated.

[11] The 1SG.DAT/1SG.GEN pronoun, *ngajuku*, involves the historical addition of the dative suffix to an earlier dative stem. Despite the transparency of this construct, the pronoun can be further inflected with the dative.

978–0–19–920647–6 53-Spencer-ch52 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 768 of 769 May 16, 2008 19:24

(24) *Marrkara-partu wurnta-la nyunanga, wurnta-la nyunanga*
brother-2SG.POSS.NOM come-PRES 2SG.DAT come-PRES 2SG.DAT
wiya-larta.
see-PURP.NOM
'Your brother is coming to you, coming to see you.'

(25) *Wiya-larta yarra-ngka jananga-mu.*
see-PURP.NOM go.ANT-2SG 3PL.DAT-ANT
'You've got to go to see them.'

Analysing apparent pronominal objects of purpose clauses as external dative arguments of the matrix clause is a necessary analysis for examples such as (25) in which the dative pronoun is within the scope of the matrix anticipatory mood construction.[12] Such an analysis immediately calls into question the status of any apparent dative object of a subordinate clause. These too might be treated instead as external dative objects of the matrix clause.[13] Like the dative, the ergative suffix cannot be followed by any further inflection. In fact, there are very few contexts in which an ergative nominal might be inflected since the ergative suffix has no adnominal functions. However, nominals denoting subjects of transitive verbs can appear in subordinate clauses, and thus there is the potential for ergative forms to bear complementizing case. But this does not happen. As (21) illustrates, apparent subjects in object relative clauses simply bear the appropriate associating case suffix preceding the complementizing suffix. This suggests that the associating case replaces the expected ergative suffix. Unfortunately, there are no examples in the existing corpus in which a second predication on a transitive subject appears in a subordinate clause with a subject pivot and so it is not possible to test whether or not ergative marking appears in this context.

The data does not allow a detailed description of the operation of the Repeated Morph Constraint (Menn and MacWhinney 1984) as it applies in Nyamal.[14] Although there are a number of contexts in which the repetition of the same suffix form might occur – albeit with different functions – tests for sequences of repeated morphemes have not been conducted.

[12] This involves a discontinuous verbal inflection, *-a-mu*, which brackets subject agreement on the verb (19) but may also embrace an immediately following accusative, dative, or locative pronoun complement of the verb (20).

[13] Dench (2006) argues that exactly this pattern, and a subsequent reanalysis of the dative NPs as arguments of the subordinate clause, provides a diachronic explanation for patterns of subordinate clause head-marking in other Pilbara languages.

[14] For a discussion of the role of this constraint with respect to multiple nominal suffixing in particular, see Moravscik (1995). Some recent accounts have attempted to incorporate the Repeated Morph Constraint into a more general Obligatory Contour Principle, see for example Yip (1998).

978–0–19–920647–6 53-Spencer-ch52 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 769 of 769 May 16, 2008 19:24

52.3 Conclusion

This chapter has presented a brief description of complex case-marking patterns in Nyamal. The different parameters identified here and allowing a description of the Nyamal system are at work in a wide range of Pama-Nyungan Australian languages, though very few show the degree of complexity found here. Nyamal is an extreme example of a dependent-marking language with few constraints on the multiple embedding, and marking, of subordinate structures – but is not representative of the majority of Australian languages.

This level of complexity begs the question of how such a system might have developed. In my view (Dench 2006), the pattern results from the maintenance of a rule of complete concord for case-marking (§52.2.2) – more typical perhaps of languages with flat, or non-configurational, structures – alongside the evolution of embedded (and thus constituent) structures through the instantiation of the argument structures of nominalized verbs.

CHAPTER 53

CASE IN AN AUSTRONESIAN LANGUAGE

DISTINGUISHING CASE FUNCTIONS IN TUKANG BESI

MARK DONOHUE

53.1 Tukang Besi

The 'classic' paradigm of case marking assumes a unique and (pragmatically, semantically, or syntactically) coherent meaning for each phonologically distinct case. Thus in the Japanese examples in (1) each of *wa*, *ga*, *de*, and *o* marks a unique role for the NP it follows, definable either pragmatically (for *wa*, 'topic'), semantically (for *de*, 'instrument' and 'location of event') or syntactically (*ga* 'nominative' and *o* 'accusative').

Japanese

(1) a. *Onna wa kooen de tori o mita.*
woman TOP park LOC bird ACC watched
'The woman watched birds in the park.'

978–0–19–920647–6 54-Spencer-ch53 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 771 of 778 May 16, 2008 19:1

b. *Tori wa onna ga kooen de mita.*
bird TOP woman NOM park LOC watched
'The woman watched birds in the park.'

A different set of organizational principles govern case in Tukang Besi, an Austronesian language of central Indonesia (Donohue 1999). There is an unproblematic genitive case, an oblique case that marks modality, and two cases that mark core arguments ('terms'), *na* and *te*.[1] With a bivalent clause such as (2) *na* and *te* are used to mark the A and the P, respectively. We can see that these case markers may not be reversed, as in (2b), and that the position of the two arguments is fixed, as shown by the ungrammaticality of (2c).

(2) a. *No-'ita te kadadi na wowine.*
she:saw TE bird NA woman
'The woman watched birds.'
b. **No-'ita na kadadi te wowine.*
she:saw NA bird TE woman
c. **No-'ita na wowine te kadadi.*
she:saw NA woman TE bird

With monovalent clauses the S is marked with *na*, as can be seen in (3).

(3) a. *No-kede na wowine.*
she:sat NA woman
'The woman sat down.'
b. **No-kede te wowine.*
she:sat TE woman

Based on the data in (2) and (3) alone we would have no trouble distinguishing *na* as a nominative case marker, and *te* as the accusative. The language is similar to Japanese in terms of the alignment of the case markers. Complications quickly arise with examples such as (4) and (5).

(4) *Te wowine no-'ita te kadadi.*
TE woman she:saw TE bird
'The woman watched birds.'

(5) *Te wowine no-kede.*
TE woman she:sat
'The woman sat down.'

[1] I describe these four as case markers, rather than prepositions or articles, since (i) certain prepositions may govern case markers, and some prepositions do not require overtly case-marked NPs, but no case-markers govern prepositions on their NPs, and (ii) the form of the case marker correlates with grammaticality and interpretation in a very fundamental way, correlating with clausal position and verbal marking, not features usually associated with articles, and the occurrence of demonstratives is independent of the choice of case marker. The fact that there are pragmatic constraints on some of the case markers implies that their history is probably intertwined with articles.

978–0–19–920647–6 54-Spencer-ch53 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 772 of 778 May 16, 2008 19:1

Based on (4) and (5) alone we would have to assume that TE was a syntactically empty determiner with the phonological form *te*, since there is no differentiation of different syntactic roles. In combination with the data in (2) and (3) we must advance a more complicated characterization of the two case markers *te* and *na*, as seen in (6).

(6) *te* marks nominative case when preverbal;
it marks accusative case when postverbal.
na marks nominative case when postverbal.

The fact that clauses of the form **te kadadi no'ita na wowine* are ungrammatical supports the hypothesis in (6): preverbal *te* does not mark accusative case. Although our characterization of case in Tukang Besi must make reference to position in the clause as well as to morphological form, it is still readily characterizable as a nominative–accusative system, with complications based on position such as have been described by, for instance, Marantz (1984). There are, however, further complications. In the examples seen so far the bivalent verbs have all shown S,A agreement by prefix (*no-* for third person arguments in realis clauses). It is also possible for the P to be cross-referenced on the verb. When there is agreement for the P *na* marks the P, and *te* the A, as in the two (equivalent) sentences in (7). By comparison with (2) we can see that this is exactly the opposite set of form:meaning correspondences as were seen there. Indeed, if we were to compare the case used in the monovalent clause in (3) with the bivalent data in (7) alone, we would have to characterize *te* as the ergative, and *na* as the absolutive case marker.[2]

(7) a. *No-'ita='e te wowine na kadadi.*
she:saw:them TE woman NA bird
'The woman watched the birds.'
b. *No'ita'e na kadadi te wowine.*

Examining the possibilities for preverbal positions in clauses with P agreement on the verb, seen in (8), and combining with the monovalent data in (3) and (5), as well as the bivalent data in (7) (but ignoring, for the moment, (2) and (4)), we would arrive at the characterization of case seen in (9) (noting in passing the ungrammaticality of **te wowine no'ita'e na kadadi*).

(8) *Te kadadi no-'ita='e te wowine.*
TE bird she:saw:them TE woman
'The woman watched the birds.'

(9) *te* marks absolutive case when preverbal;
it marks ergative case when postverbal.
na marks absolutive case when postverbal.

[2] An analysis that treats clauses such as (7) as truly bivalent, and clauses such as (4) as 'antipassive' or 'detransitive' variants of these is untenable. See Donohue (1999: 158–66) for discussion.

In fact, of course, we cannot ignore the existence of the data in (2) and (4). Although the clauses with P-agreement present a clear increase in the amount of morphology in the clause, they are more frequent in natural speech, and P-agreement markers are acquired earlier than the nominative agreement prefixes. Following the argumentation in Donohue (1999: chapter 7), I propose that both clause types, both with and without P agreement markers, are basic.

There is clearly no one-to-one correspondence between the case marker used and the syntactic role of an argument, in terms of argument-structure or semantic features that will neatly correspond to categories such as 'nominative' and 'accusative' in Japanese, or 'ergative' and 'absolutive' in other languages. We can, nonetheless, offer some robust characterizations of the case system, shown in (10).

(10) *te* marks a preverbal argument which would have been eligible to receive *na* case if postverbal;
it marks a postverbal argument that is not eligible to appear preverbally.
na marks a P when it is postverbal and shows agreement on the verb; if there is no agreement on the verb for P, *na* marks the A (or S, if monovalent).

In terms of directly coding the arguments of the clause, the reader would be excused for thinking that the Tukang Besi case-marking system, in contrast to the rather transparent system of verbal agreement, is convoluted. This depends on our opinions on what a case-marking system is intended to encode. There are different opinions on this, and the widely accepted distinction between structural and semantic cases reflects this dichotomy. Examining the more complex case systems (e.g. Daniel and Ganenkov, Chapter 46) we would have to conclude that case exists to mark semantic distinctions. In the more simple case systems (e.g. König, Chapter 50, or Foley 1991 on Yimas) the relationships marked cannot be easily described in terms of semantic distinctions, but exist in terms of syntactic distinctions. Just as a 'standard' nominative–accusative case-marking system does not mark semantic roles, but more directly grammatical information, so too can the Tukang Besi case system be more productively viewed not as encoding specific information about the identity of the A and P, but rather about the SUBJECT and the OBJECT.[3] The full case system of Tukang Besi consists of four members, shown in (11) (there are also a number of prepositions, some of which subcategorize for a *te*-marked object).

(11)

na	subject
te	object
i ∼ *di*	oblique / adjunct (unmarked/irrealis ∼ past tense/realis)
nu	genitive

[3] The question of grammatical functions in western Austronesian languages has a long history, though it was only brought to the attention of general linguists by Schachter (1976, 1977). Work in a variety of frameworks by Guilfoyle (1992), Kroeger (1993), Sells (2000) et al., and Pearson (2005), demonstrates that the issues are not as problematic as presented by Schachter's articles.

978–0–19–920647–6 54-Spencer-ch53 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 774 of 778 May 16, 2008 19:1

When we examine syntactic tests for grammatical functions (see, for instance, Dalrymple 2000 for cross-linguistically relevant constructions) we find that the *na*-marked argument is consistently the most privileged argument of the clause – in other words, the subject (though see Donohue 2005b). To present just one example, examine (12)–(14), which illustrate the scope of floating quantifiers in different clauses. In an intransitive clause the scope of *saba'ane* is unambiguously with the argument, not with the adjunct *di wunua*. In each of (13) and (14) there are two arguments, and *saba'ane* scopes over the *na*-marked argument, regardless of whether it is the A or the P.

(12) *Saba'ane no-moturu di wunua na ana.*
all they:sleep DI house NA child
'All the children slept in the house(s).'
* 'The children slept in all the houses.'

(13) *Saba'ane no-'ita te 'obu na beka.*
all they:saw TE dog NA cat
'All the cats saw the dogs.'
* 'The cats saw all the dogs.'

(14) *Saba'ane no-'ita='e te beka na 'obu.*
all they:saw:them TE cat NA dog
'The cats saw all the dogs.'
* 'All the cats saw the dogs.'

Further arguments for the privileged status of the *na*-marked argument can be found in Donohue (1999, 2004), and shall not be repeated here. There is considerable debate about whether the alternation in (13) and (14) (and earlier (2) and (7)) represents a true alternation in the status of the subject or not in Austronesian languages (e.g. Aldridge 2004, Pearson 2005), but whatever the exact syntactic status of the alternation, it is clear that there is a change in syntactic privileges when the cross-referencing on the verb changes.

Another factor that is relevant to our differentiation of the case markers *te* and *na* is pragmatic. In (15) and (16) we can see that, while there are no restrictions on questioning arguments *in situ* (there are also pseudo-cleft strategies), the *na*-marked argument may not be questioned.[4] Furthermore, preverbal position and the *te* case marker does not license questioning: the same restrictions that were found with postverbal arguments are found with preverbal ones, as shown in (17) (compare with (15b) and (16a)).

(15) a. *No-'ita te paira na wowine?*
she:saw TE what NA woman
'What did the woman see?'

[4] This means that the pseudo-cleft strategy is the only one available for questioning the subject of a monovalent clause.

978–0–19–920647–6 54-Spencer-ch53 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 775 of 778 May 16, 2008 19:1

b. **No-'ita te kadadi na emai?*
she:saw TE bird NA who
'Who saw the bird?'

(16) a. **No-'ita='e te wowine na paira?*
she:saw:it TE woman NA what
'What did the woman see?'
b. *No-'ita='e te emai na kadadi?*
she:saw:it TE who NA bird
'Who saw the bird?'

(17) a. **Te emai no-'ita te kadadi?*
TE who she:saw TE bird
'Who saw the bird?'
b. **Te paira no-'ita='e te wowine?*
TE what she:saw:it TE woman
'What did the woman see?'

What, then, is the status of arguments in the preverbal position? We have seen that only the argument that we would expect to see marked with *na* may appear there, and that in that position the argument is marked by *te*. The difference between (2) and (4), and between (7) and (8), is that in (4) and (8) a degree of identificational focus is necessarily associated with the preverbal argument; this identificational focus cannot be found with the equivalent postverbal argument, as shown in (18), which presents responses to a question calling for identification of the subject.

[Question: Who went home?]

(18) a. *Te wowine no-mbule=mo.*
TE woman she:returned
'The woman went (home).'
b. *#No-mbule=mo na wowine.*
she:returned NA woman

How do we characterize the alignment of case marking? Both (2a) and (7a) can be argued to be 'basic' clauses in the language, and so both must be considered when determining alignment – and unfortunately our notion of alignment does not extend to there being two basic transitive clause types. Rather than being characterized as showing nominative–accusative or ergative–absolutive or semantic alignment, the case systems of the western and northern Austronesian languages directly mark grammatical functions, bypassing the standard notions of 'alignment' entirely. Even then, we need a set of conventions to interpret the 'meaning' of the case markers, as shown in (19).

(19) *te*: if preverbal, *te* marks the subject;
if postverbal, *te* marks a core non-subject ('object').
na: necessarily postverbal, *na* marks the subject.

Examined in terms of pragmatic information, we can re-couch the generalization in (19) as (20).

(20) *te*: if preverbal, *te* marks identificational focus (see (18));
if postverbal, *te* marks a non-given term.
na: necessarily postverbal, *na* marks a (non-focused) given term.

The complexities do not end here. Since first or second persons are necessarily given information in any speech act, we would predict that it is impossible for *te* to mark a postverbal argument if it is first or second person. To some extent this prediction is borne out, as can be seen in (21). A 1SG argument may appear preverbally with *te*-marking, but not postverbally. For a third person argument these restrictions do not hold.

(21) a. *Te iaku no-'ita=aku.*
TE 1SG they:saw:me
'They saw me (and not someone else).'
b. *#!No-'ita te iaku.*
they:saw TE 1SG
'They saw me.'

(22) a. *Te ia no-'ita='e.*
TE 3SG they:saw:her
'They saw her (and not someone else).'
b. *No-'ita te ia.*
they:saw TE 3SG
'They saw her.'

It is, however, possible for *te* to mark a first or second person argument when it is postverbal, but only when that argument is the A of a clause such as (7). Reversing the roles of the participants of (21) results in the clauses in (23), both of which are grammatical despite the presence of a postverbal *te* phrase.

(23) a. *Te ia ku-'ita='e (te iaku).*
TE 3SG I:saw:her TE 1SG
'I saw her (and not someone else).'
b. *Ku-'ita='e te iaku (na ia).*
I:saw:her TE 1SG NA 3SG
'I saw her.'

The Tukang Besi case system, then, does depend on syntactic roles to some extent, despite being overwhelmingly governed directly by grammatical function information, rather than argument-structure information.

978–0–19–920647–6 54-Spencer-ch53 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 777 of 778 May 16, 2008 19:1

53.2 Other western Austronesian languages

The pattern of case marking found in Tukang Besi, directly marking grammatical functions rather than marking specific positions in argument structure, is one that is shared by many languages of the so-called 'Philippine' type (see Donohue 2002 for a discussion of the reflexes of the Philippine voice morphology in Tukang Besi). The sentences in (24) (from Tagalog) show the same alternation of case marking between the A and the P that we observed in Tukang Besi between (2) and (7). As in Tukang Besi, the variation in *ang* marking the A or the P corresponds to the affixation on the verb. While in Tagalog the verbal affixes do not show a paradigm of inflection as the Tukang Besi cross-referencing does (though see Sells 1998, 2000), the net effect is the same.

Tagalog

(24) a. *Nakita ng bata ang aso.*
PV:saw GEN child NOM dog
'The child saw the dog.'
b. *Nakakita ng aso ang bata.*
AV:saw GEN dog NOM child
'The child saw a dog.'

Note that the case marking associated with the nominative argument does not vary by position in the clause, unlike the Tukang Besi examples seen earlier. (25) shows that Tagalog, which represents a more typical western Austronesian case-marking system, marks both preverbal and postverbal subjects with *ang*. Similarly, rather than having a dedicated core case marker, the genitive *ng* is used for non-*ang* arguments. The attractions of a nominal analysis of the data that are engendered by such a syncretism (under which (24a) would be more directly translated as 'The dog is the seen-one of the child') have been pointed out by many authors, just as many others have pointed out the problems with such an analysis.

(25) a. *Ang aso-(a)y nakita ng bata.*
NOM dog-AY PV:saw GEN child
'The child saw the dog.'
b. *Ang bata-(a)y nakakita ng aso.*
NOM child-AY AV:saw GEN dog
'The child saw a dog.'

The Austronesian patterns seen here, primarily through Tukang Besi, reveal direct reference to something other than argument structure positions. This chapter began with a brief look at Japanese, which has a pragmatic NP-marking system

978–0–19–920647–6 54-Spencer-ch53 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 778 of 778 May 16, 2008 19:1

separate from, though integrated with, the nominative–accusative syntactic system. In (western) Austronesian case systems too, pragmatic information plays a part, but it is completely integrated with syntactic marking. Since grammatical notions such as 'subject' are, it has been argued, the syntacticization of discourse-salient positions, this should not be seen as surprising. The fact that such tight integration of pragmatics and syntax is rare around the world does not mean that these case-marking systems do not deserve deeper investigation.

CHAPTER 54

CASE IN A TOPIC-PROMINENT LANGUAGE

PRAGMATIC AND SYNTACTIC FUNCTIONS OF CASES IN JAPANESE

AKIO OGAWA

54.1 Characteristics of topic-prominent languages

Discussions about topic-prominent vs. subject-prominent languages are old but still of current interest. According to the now classic work of Li and Thompson (1976), topic-prominent languages possess the following characteristics: (a) the topic is coded on the surface, that is morphologically or/and syntactically; (b) passive constructions either do not or only marginally exist or carry a special meaning; (c) there are no dummy or empty subjects; (d) double subject constructions are available; (e) it is not the subject, but the topic that controls coreferential constituent deletion; (f) verb-final languages tend to be topic-prominent; (g) there

978–0–19–920647–6 55-Spencer-ch54 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 780 of 788 May 16, 2008 19:1

are no constraints on what kind of constituent may be the topic; and (h) topic–comment sentences are basic.

Japanese actually fulfils these criteria more or less: (a) it has the topic marker ∼*wa*; (b) there are adversative passives that are fundamentally not based on the active–passive diathesis and carry a special meaning (cf. Palmer 1994, Shibatani 1994); (c) meteorological or chronological expressions, for which subject-prominent languages typically employ a dummy subject, exhibit a referential subject or no subject at all (cf. Ogawa 2006); the characteristics (d), (e), and (f) fully apply, whereas (g) and (h) apply either to a certain extent or need more evidence and argument.

First, the topic-prominent character of Japanese will be discussed with special reference to its case system. I will focus on Li and Thompson's criteria (a), (d), and (g). After sketching synchronic and diachronic case-drop phenomena, some other typologically related as well as unrelated languages will be briefly exemplified for their functional parallelism.

54.2 Case system of Japanese

The traditional grammar of Japanese distinguishes between two types of postpositional particles: *kaku-joshi* (case particles) on the one hand and *kakari-joshi* (relating/charging particles) on the other hand. The latter historically needed a correlated verbal form (a sort of agreement) in the sentence final position (e.g. ∼wa . . . + assertive verbal from), which is no longer true in Modern Japanese.

Japanese makes use of a number of postpositional particles in the case system such as: sensei∼*ga* (teacher-nominative), ∼*wo* (accusative), ∼*ni* (dative), ∼*no* (genitive), ∼*de* (locative, instrumental), ∼*e* (directional), ∼*kara* (ablative), ∼*to* (comitative). These 'kaku-joshi' occur in almost perfect complementarity with each other within a sentence and do not violate the 'single case condition', albeit in some cases (e.g. ∼*ni* vs. ∼*de*, ∼*ni* vs. ∼*e*) there are syntactic or semantic overlaps (cf. Yamanashi 1995).

While all of the above-mentioned postpositions mark cases, i.e. grammatical relations, the topic marker ∼*wa* which belongs to 'kakari-joshi' assigns no case as such and roughly means 'as for X'. Apart from marking a grammatical relation, the particle ∼*mo* (in the sense of 'also', 'even') also plays an important role. However, it will be not discussed here, since the topic-prominent character of Japanese is sufficiently obvious from the behaviour of ∼*wa*, especially compared with the ∼*ga* case marking.

978–0–19–920647–6 55-Spencer-ch54 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 781 of 788 May 16, 2008 19:1

As for generative treatment of the case system in Japanese I will not address it in this article, but see Inoue 2006 for an extensive discussion.

54.3 Pragmatic and syntactic functions of ~*wa* and related constructions

54.3.1 ~*wa* and ~*ga*

There are to some extent controversial opinions about the function(s) of ~*wa*. Although Iwasaki (1987) sees the core function of ~*wa* as 'scope setter' from which its topic function is merely derived, and although Kuroda (2005) even argues against ~*wa* as a topic marker (in a strict sense), it undoubtedly determines the topic-prominent character of Japanese.

The topicalizing function of ~*wa* can most clearly be shown in comparison with ~*ga*, because ~*wa* and ~*ga* can be considered as counterparts in many respects:

(1) a. *Kare-wa Yamada-san desu.*
he-TOP Yamada-Mr./Ms. be
'He is Mr. Yamada.'
b. *Kare-ga Yamada-san desu.*
he-NOM Yamada-Mr./Ms. be
'Mr. Yamada is he.'

The nominal marked with ~*wa* is the topic, the predicate nominal (plus the copula) the comment, as in (1a). In (1b), on the contrary, the predicate nominal is the topic, and the nominal followed by ~*ga* is the comment. Here terms such as topic and comment are used in reference to the information structure status of a constituent.

There have been ample discussions about the distinction between ~*wa* and ~*ga* for the information structure (cf. Chafe's 1971 differentiation between 'old information' and 'new information', Kuno's 1973 survey on this theme, and a lot of works published in Japanese). In this sense ~*ga* functions not only as a nominative marker but also as an 'anti-topic' one.

It is predictable that ~*wa* can rarely be employed in subordinated sentences. For they essentially contradict the topic–comment structure. The subject marker ~*ga*, on the other hand, occurs also in embedded sentences.

(2) *Kare-ga/*-wa kuru-node, . . .*
he-NOM/-TOP come-because
'Because he will come, . . .'

978–0–19–920647–6 55-Spencer-ch54 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 782 of 788 May 16, 2008 19:1

~wa in its topicalizing function can lead to strongly 'contrastive' interpretations like in:

(3) *Kare-wa eigo-wa hanas-eru.*
he-TOP English-TOP speak-can
'He speaks English.'

The sentence-initial referent 'kare' is already topicalized; the following one 'eigo', also marked with *~wa*, is furthermore topicalized, being so to speak 'picked up' from other related or evocable ones. In (3) there hence occurs a strongly contrastive effect: 'He can speak no other foreign language but English.'

~ga can function as an 'exclusion' such as:

(4) *Kono shigoto-wa watashi-ga yarimasu.*
this task-TOP I-NOM do-DIS
'I will do this task.'

This sentence can be paraphrased like this: 'Nobody else but me will do this task'. So persons other than 'me' are excluded. 'Watashi', the speaker, being attributed high topicality per se, essentially becomes a 'comment' through *~ga*, resulting in a kind of mirror effect.

It seems that 'contrast' by *~wa* on the one hand and 'exclusion' by *~ga* on the other hand lead to similar – albeit contradictory – interpretations. (Such difficulties in differentiation between *~wa* and *~ga* motivate Kuroda 2005 not to regard *~wa* as a topic marker.) The former, however, is based on the process in which one stresses the topic, i.e. old (already-provided) information or information that the speaker supposes the hearer to know (cf. Onoe 1995). The latter results from the process in which one emphasizes the comment to a still higher degree, that is, the speaker presents to the hearer some information as entirely new, unknown at all, even if the item marked by *ga* apparently is not, such as 'watashi' in (4). Both processes are commonly 'picking up', but from different perspectives, namely foregrounding of the topic vs. of the comment.

54.3.2 Word order

Japanese reveals a relative freedom of word order, albeit it exhibits a rigid verb-final position. There are, however, some canonical word order patterns, to which the 'nominative–dative–accusative' belongs:

(5) *Sensei-ga seito-ni hon-wo ageta.*
teacher-NOM pupil-DAT book-ACC give-PAST
'The teacher gave the pupil a book.'

As for topicalization, two strategies are available: i) the use of the topic marker *~wa*, as we have already seen, and ii) the left-dislocation of the nominal with its

case-marking unchanged. The former (frequently together with the latter, unless the nominal in question is the subject) leads to topicalization of the nominal:

(6) *Sono-hon-wa sensei-ga seito-ni ageta.*
this-book-TOP teacher-NOM pupil-DAT give-PAST
'As for this book, the teacher gave it to the pupil.'

The latter lets the nominal in question become a comment, hence the rest is a topic in the information structure:

(7) *Sono-hon-wo sensei-ga/-wa seito-ni ageta.*
this-book-ACC teacher-NOM/-TOP pupil-DAT give-PAST
'It is this book that the teacher gave to the pupil.'

This kind of left-dislocation, that is, not the topicalization through ∼*wa*, often results in an 'exclusive' interpretation, as we have seen in the case of ∼*ga* in (4). While ∼*wa* marks no grammatical relation as such, ∼*ga* is a nominative/subject marker, so it cannot be used for the accusative nominal in (7).

As in (6), the topicalization through ∼*wa* can be applied not only to genuine subject nominals, but also to object nominals:

(8) a. *Taro-wa Hanako-wo aishiteiru.* (originally nominative *-ga*)
Taro-TOP Hanako-ACC love
'Taro loves Hanako.'

b. *Hanako-wa Taro-ga aishiteiru.* (originally accusative *-wo*)
Hanako-TOP Taro-NOM love
'It is Taro who loves Hanako.'

c. *Hanako-wa Taro-ga denwashita.* (originally dative *-ni*)
Hanako-TOP Taro-NOM phone-PAST
'It is Taro who phoned Hanako.'

On the contrary, oblique nominals can hardly ever be topicalized by means of ∼*wa*, such as:

(8) d. **Hanako-wa Taro-ga Tokyo-e itta.* (originally comitative *-to*)
Hanako-TOP Taro-NOM Tokyo-DIR go-PAST
'It is Taro who went to Tokyo with Hanako.'

e. **Naifu-wa Taro-ga pan-wo kitta.*
knife-TOP Taro-NOM bread-ACC cut-PAST
'With the knife Taro cut the bread.'
(originally locative-instrumental *-de*)

But as long as the original case marker remains, a more argument-like (less oblique) nominal can be topicalized by attaching ∼*wa* (cf. [8d]: '*Hanako-to-wa*' is well-formed), a less argument-like (more oblique) one even then can not (cf. [8e]: '*Naifu-de-wa*' is less possible).

With regard to the following statement: 'in topic-prominent language [. . .] there are no constraints on what may be the topic' (Li and Thompson 1976: 471), Japanese actually might be characterized as a strongly, but not fully topic-prominent language. In this respect the data under (8) suggest a parallelism to the Noun Phrase Accessibility Hierarchy (Keenan and Comrie 1977), according to which relativization follows the hierarchy within the case system; relativization is one of the strategies for topicalization.

54.3.3 Predication and double subject constructions

In the case of ∼*wa* an extremely broad range of predicative relations can be attested. Some of them are:

(9) a. *Kare-wa Yamada-san desu.*
he-TOP Yamada-Mr./Ms. be
'He is Mr. Yamada.'
b. *Kare-wa byouki desu.*
he-TOP ill be
'He is ill.'
c. *Kare-wa koohii desu.*
he-TOP coffee be
'He ordered coffee.'
d. *Kare-wa onnanoko desu.*
he-TOP girl be
'He has a daughter.'

Subject-prominent languages, on the contrary, allow only a few of these uses – (9a) ('identification') and (9b) ('property') – provided in copular constructions. Though in English for example one can say 'You are a coffee, right?' as spoken by a waitress in a restaurant. But the range of possibilities of such expressions is much smaller than in Japanese.

Ikegami (1981) characterizes the function of this copular (or copular-like) construction in Japanese as 'A WITH B'. This can be interpreted as: A is in some relation with B. If, for example, 'he' orders 'coffee' in a restaurant, 'he' is in some relation with 'coffee'. The sentence (9c) hence allows not only the interpretation glossed above, but also interpretations like: 'He dislikes coffee', 'He plants coffee', 'He sells coffees', etc. Which one is suitable can only be verified pragmatically.

To add: The copular-like construction can employ ∼*ga* too, instead of ∼*wa*. This results in a mirror relation in terms of the topic–comment structure, as seen in (7).

The function of ∼*wa* can also be compared to 'opening the file' concerning the category about which an assertion should be made. The operation of 'opening the file' can typically be applied to so-called double subject constructions, such as:

978–0–19–920647–6 55-Spencer-ch54 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 785 of 788 May 16, 2008 19:1

(10) *Zou-wa hana-ga nagai.*
elephant-TOP trunk-NOM long
'An elephant has a long trunk.'

(11) *Nihon-wa rokugatsu-ga ichban ame-ga ooi.*
Japan-TOP June-NOM most rain-NOM much
'In Japan it rains most in June.'

In (10) the file of the category 'elephant' is opened first. It evokes a number of 'facets' concerning the 'elephant'; one of them is 'the trunk is long' (cf. Sawada 2002). (11) is a 'triple' subject construction, which is coherent with the idea of opening the file for multiple applications. The characterization of 'staging' through ~*wa* (Maynard 1980) also supports this idea. And for the process 'opening the file' or 'staging', the 'aboutness condition' (Shibatani 1994) is relevant.

Some double subject constructions semantically overlap with genitive constructions:

(12) a. *Zou-wa hana-ga nagai.*
elephant-TOP trunk-NOM long
'An elephant has a long trunk.'
b. *Zou-no hana-ga/-wa nagai.*
elephant-GEN trunk-NOM/-TOP long
'The trunk of an elephant is long.'

These two construction types, however, fundamentally differ. (12a) is an assertion about an elephant, whereas (12b) is an assertion about an elephant's trunk. This does not seem to be a case of possessor raising, but a kind of expansion of valency of a predicate (cf. Shibatani 1994, Ogawa 1997). Other double subject constructions, quite typical ones, cannot be paraphrased with genitive constructions:

(13) a. *Watashi-wa atama-ga itai.*
I-TOP head-NOM hurt
'I have a headache.'
b. **Watashi-no atama-ga itai.*
I-GEN head-NOM hurt
'My head hurts.'

While genitive constructions must be adjacent, double subject constructions need not (cf. Heycock 1993):

(14) a. *Zou-wa umaretsuki hana-ga nagai.*
elephant-TOP from.the.birth trunk-NOM long
'An elephant has a long trunk from birth.'
b. **Zou-no umaretsuki hana-ga/-wa nagai.*
elephant-GEN from.the.birth trunk-NOM/-TOP long
'The trunk of an elephant is long from birth.'

978–0–19–920647–6 55-Spencer-ch54 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 786 of 788 May 16, 2008 19:1

The possibility of discontinuity indicates the double subject as being a sentential constituent.

54.3.4 'Case drop' synchronic and diachronic

The case marker can drop especially in colloquial speech, as long as its grammatical relation is reconstructable:

(15) *Aitsu, mou ronbun kaita?*
he-ϕ already article-ϕ write.PAST
'Has he already written his article?'

Here the nominative argument *aitsu* as well as the accusative one *ronbun* respectively lack their case markers.

In general, case markers that are highly argument-like, i.e. nominative and accusative, can easily drop. Oblique case markers, however, can drop too, provided the argument in question is – mostly idiomatically – incorporated into the verb:

(16) a. *Sonoatode gakko itta-yo.*
after.that school-ϕ go.PAST-DIS
'After that I went to school.'
b. *Mou ie kaeroo-ka?*
yet house-ϕ go.back-DIS
'Shall we go home?'

In terms of topic prominence the comparison between *∼wa*, *∼ga*, and case-dropping (see below) is crucial (cf. Shibatani 1990, Frey 2003):

(17) a. *Watashi-wa samishii-no.*
I-TOP sad-DIS
'I am sad.'
b. *Watashi-ga samishii-no.*
I-NOM sad-DIS
'It is I who is sad.'
c. *Watashi samishii-no.*
I-ϕ sad-DIS
'I am sad.'

Watashi-wa in sentence (17a) is strongly topicalized. The dichotomy of the topic–comment structure is in full use. The counterpart lies in (17b), where *watashi-ga* is the comment. The sentence (17c) without overt case-marking can be located between them. The topic–comment structure is neutralized, *watashi* can be seen as 'quasi-topic'. A 'deictic/stage-dependent' (Onoe 1995) argument, *watashi* representatively, possesses such high topicality that it tends to occur without *∼wa*. The

sentence (17c) hence is an unmarked utterance, while (17a) and (17b) are topic–comment structurally marked. We rather should speak not of the dropping of ∼*wa* but of its non-existence.

Historically ∼*wa* triggered a sentence-final particle, which is known as *kakari-musubi* ('charging-closing'). Recall: ∼*wa* belongs to 'kakari-joshi' ('charging/relating particles') in the traditional grammar of Japanese. The 'kakari-musubi' is again confirmation of the theoretical linguistic claim that the main function of ∼*wa* lies in scope setting (Iwasaki 1987): the domain must have been set up between both the co-occuring particles. In the era of 'kakari-musubi' the kakari ('charging') particle ∼*wa* assigned the beginning of the scope, and a corresponding *musubi* ('closing') particle overtly marked its end. Besides 'kakari-musubi' there existed (and – as we have seen in this section – partly still exists) zero case-marking, especially in subject as well as object position, whereas the topic function attributed to ∼*wa* in Modern Japanese was broadly borne by zero marking (Kinsui 1995). The 'kakari-musubi' itself progressively vanished in the Kamakura era (twelfth–fourteenth century).

This correlated with ∼*ga* obtaining the status of subject in the Kamakura era and the following Muromachi era (fourteenth–sixteenth century). ∼*ga* had persisted till the Heian era (eleventh century) as the genitive marker. A remnant of this can still be observed in its overlapping use with ∼*no* (*Taro-ga/-no katta hon*: Taro-NOM/-GEN buy-PAST book).

54.4 Parallel data in some other languages

There are a number of grammatical strategies that distinguish topic and comment. Some languages employ their topic marker quite often, such as Japanese and slightly differently Korean or Burmese:

(18) Korean:
Ku saram-un/-i Park imnida.
this person-TOP/-NOM Park be
'He is Mr. Park.'

(19) Burmese:
e:dilou pyo:da ha məkaun:bu:no.
such a thing to.say-TOP bad-DIS
'It is bad to say such a thing.' (Kobayashi 1984: 93)

978–0–19–920647–6 55-Spencer-ch54 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 788 of 788 May 16, 2008 19:1

For functional equivalence other languages use their article system and/or word order variations and/or different sentence types. For example:

(20) German:
a. *Der Mann heiratet die Frau.*
the man-NOM marry the woman-ACC
'The man marries the woman.'
b. *Die Frau heiratet der Mann.*
the woman-ACC marry the man-NOM
'It is the man who marries the woman.'

(21) Spanish:
La tortilla la comió Pedro.
The omelette-ACC/-NOM it-ACC eat-PAST Pedro
'As for the omelette, Pedro ate it.' (Noda 1996: 294)

(22) Chinese:
a. *Lao-Wáng zuótiān xiū le zhe liang zìxíngchē.*
Mr. Wang yesterday fix PRF this bicycle
'Mr. Wang fixed this bicycle yesterday.'
b. *Zhe liang zìxíngchē shi Lao-Wáng zuótiān xiū de.*
This bicycle it Mr. Wang yesterday fix DIS
'As for this bicycle Mr. Wang fixed it yesterday.'

Example (20) uses the word order strategy with sentence initial position; (21) shows left-dislocation, followed by an anaphoric pronoun. (22) employs the dislocation in a different sentence type; in (22) not a particle such as the Japanese ∼*wa* but an anaphoric element (*shi*) is inserted, and corresponds to the sentence-final marker *de*. This recalls that in old Japanese ∼*wa* also required such a correlated verbal agreement in the sentence-final position (see section 54.3.4). An anaphoric element occurs in the Spanish example (21) too. In addition to these syntactic strategies for topicalization, prosodic ones play a role in all these examples.

To what extent a topic construction is more or less marked/unmarked than its counterpart or related construction can be clarified only relatively. As for Japanese, however, the topic marker alternates with the argument-like case markers (nominative, accusative, dative) and can also co-occur with the rest (oblique case markers). The bare left-dislocation of grammatically case-marked nominals also results in a topic–comment structure. Both undoubtedly indicate an unmarked base for a topic-prominent language.

CHAPTER 55

CASE IN YUKAGHIR LANGUAGES

ELENA MASLOVA

55.1 Introduction

This paper describes case marking of core arguments in two extant Yukaghir languages, the Tundra Yukaghir language spoken in the Lower-Kolyma region of Saha (Russia) and the Kolyma Yukaghir language of the Upper-Kolyma region. The description focuses on Tundra Yukaghir, but the relevant differences between the two languages are summarized in section 55.4.

The Yukaghir languages have rich, predominantly agglutinating morphology, which serves as the major means of expressing syntactic information. The canonical word order is head-final, but it is very flexible at the clause level and cannot be used as a guide for discrimination of core participants (see section 55.3 for a specific illustration). With minor exceptions irrelevant in the present context, verbs fall into two grammatical classes, intransitive and transitive, characterized by different inflectional paradigms. Intransitive verbs have a single core argument (S), which controls verb agreement and switch-reference in non-finite clauses. Transitive verbs have two core arguments (A and P), of which only A controls verb agreement and switch-reference. The core arguments formally differ from peripheral nominal constituents in that they can be represented by morphologically unmarked NPs (the goal/addressee arguments of ditransitive verbs are always case-marked, most commonly by dative). There is a highly grammaticalized system of expressing

978–0–19–920647–6 56-Spencer-ch55 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 790 of 796 May 16, 2008 19:1

focus structure, tightly intertwined with case marking of core participants (see section 55.2); the correspondence between the (core) argument structure and the focus structure is consistently marked on the verb (the relevant verb endings are glossed as 'AF', 'PF', and 'SF' in the examples below, where 'XF' stands for X-focus; the absence of such a label in the interlinear gloss indicates a construction without nominal focus). The case marking system is also sensitive to person hierarchy (section 55.2) and a cross-linguistically unusual grammatical classification of lexical NPs (section 55.3).

55.2 Case, person hierarchy, and information focus

The basic mechanism of discrimination between core arguments in Yukaghir is nominative–accusative, as shown by examples in (1), where A and S are unmarked, and P receives an overt marker, *-le(ŋ)*:[1]

(1) a. *qad'ir apanala: me-kelu-j*
now old.woman AFF-come-INTR(3)
'The old woman came.'
b. *qad'ir taŋ ile-le met könn'e-pul men'-ŋu-te-m titte-l'uol*
now that deer-P my relative-PL take-PL-FUT-TR(3) their-TRNSF
'My relatives will take those deer for themselves.'

The suffix *-le(ŋ)* is glossed as 'P' in (1b); this is correct, insofar as this suffix marks P whenever it occurs in a transitive clause, that is, it is incompatible with A. It does not mean, however, that this is a canonical accusative case. To begin with, its presence depends on the clause-level context; more specifically, a third-person P can remain unmarked in the context of first or second person A, as in (2).

(2) *met amā me-pun'-me-k*
my father AFF-KILL-TR-2SG
'You (SG) have killed my father.'

The lack of overt P-marking in this context never leads to ambiguity, because the person of A is marked on the verb, so an unmarked third person NP can only be interpreted as P. If both A and P are first or second person, P takes a special accusative suffix, *-ul*:

[1] These and most other examples in this paper come from authentic Tundra Yukaghir texts published in Maslova (2001); sources for other examples are given after idiomatic translations.

(3) *met tet-ul ja-n sukunmol'hil-han ura-rit'i-t*
I you-P three-AT year-PROL learn-CAUS-FUT(TR.1SG)
'I will be teaching you for three years.' (Krejnovič 1958: 57)

Thus, the syntactic role of first and second person pronouns is always marked locally, whereas the role of an unmarked third person NP can only be established on the basis of its clause-level context.

A further complication arises with examples such as (4) and (5):

(4) *met ten'i n'awn'iklie-leŋ toŋore-meŋ*
I here polar.fox-P|F chase-PF.1|2SG
'I have been chasing A POLAR FOX here.'

(5) *qahime-leŋ kelu-l*
raven-P|F came-SF
'A RAVEN came.'

In (4), *-le(ŋ)* is attached to P, even though it is not needed for P-marking (since A is the speaker); in (5), it is attached to S, in an obvious contradiction to its characterization as an accusative. In such sentences, *-le(ŋ)* functions as a marker of grammatical focus, that is, it explicitly includes the NP into the focus part of the information structure (in terms of Lambrecht 1994: 221–32); the gloss of *-le(ŋ)* is therefore replaced by 'P|F' in all examples below (where 'I' stands for (inclusive) 'or'). The F-marking is obligatory if the NP is the only focal constituent (e.g. in question–answer pairs) and optional otherwise (for a more detailed description of focus constructions in Yukaghir, see Maslova 1997; 2005).

In intransitive clauses, *-le(ŋ)* serves as a nominal F-marker; in transitive clauses, its primary function is P-marking (although the F-marking function is visible in the context of first or second person A as in (4)): in A-focus sentences, it cannot mark A, and can only appear in its P-marking function, as in (6).

(6) *nime-le aq pajp wie-nun*
dwelling-P|F only woman make-HAB(AF)
'Only WOMEN install dwellings.' (Krejnovič 1982: 210)

This case-marking pattern can be classified as a special type of 'split intransitivity' (Silverstein 1976; Dixon 1994: 70–110), where the split is determined by the locus of S in the information structure: a topical S is encoded as A, and an in-focus S, as P; Maslova 2006).[2]

[2] A possible argument against this analysis might be that the focal S is not a core argument, but rather a clause-external constituent; however, language-internal evidence indicates that the focal S is highly integrated into its clause and retains all applicable 'subject properties'.

55.3 Case paradigms and classes of NPs

The focus-oriented split-intransitivity pattern is most transparent in the examples presented above, since a single marker is used both for F and for non-focal P. This is not always the case: as shown in (7), a non-focal P can also be marked by one of the locative case markers, *-hane*∼*-qane*:

(7) *qad'ir me-lāme-s-ŋa tuŋ tolo-n solqa-hane*
NOW AFF-dog-CAUS-PL.TR(3) this wild.deer-AT crowd-LOC
'They sent a dog into this herd of wild deer.'[3]

The locative P-marking is used in the context of third person A if the NP is not eligible to receive *-le*(*ŋ*) in the F position, the eligibility being conditioned by its own inherent properties. To begin with, third person personal pronouns, proper names, and NPs with possessive modifiers cannot receive overt F-markers, that is, their F-status is marked only on the verb:

(8) *mit abut'ie joj-ā-l*
our grandmother sick-INGR-SF
'OUR GRANDMOTHER fell ill.'

The second class of NPs not eligible to receive *-le*(*ŋ*) falls into two quite distinct subclasses. First, the first and second person pronouns take another suffix, *-*(*e*)*k*, in the F position:

(9) *tet-ek n'ie-me-le*
2SG-P|F call-TR-PF.3
'She is calling YOU (not anyone else).'

In contrast to *-le*(*ŋ*), *-*(*e*)*k* cannot be used for non-focal P-marking. Accordingly, these pronouns take the locative marker as their non-focal P-marker in the context of third person A:

(10) *t'ārt'eqān mit-qane me-pun'i-m*
Charchehan we-LOC AFF-kill-TR(3).
'Charchehan has killed us!'

Secondly, *-*(*e*)*k* is used as F-marker for lexical NPs containing qualitative or quantitative modifiers (but no possessive modifiers). If the modifier constitutes the actual narrow focus of the message, then this marking is obligatory. This case is illustrated in (11): the context of this sentence suggests that it has a pragmatic presupposition like 'they sent [X] bird', and its focus contains only the name of the bird (encoded as a relative clause with copula):

[3] The causative suffix is used to derive regular monotransitive verbs from nominal stems, with a general meaning like 'cause Y to have X/be with X'.

978–0–19–920647–6 56-Spencer-ch55 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 793 of 796 May 16, 2008 19:1

(11) *Djoŋi-ŋo-d'e n'olel-ek pot'esej-ŋu-mle*
D.-be-REL bird-P|F send-3SG-PF.3
'[The birds decided to send one of them to look for a better place.] They sent a bird WHOSE NAME WAS DYONGI.'

Although the first and second person pronouns, on the one hand, and lexical NPs with focal non-possessive modifiers, on the other, do not seem to form a coherent class, they are treated alike by the Yukaghir grammar in that *-(e)k* is required in F-positions.

Apart from these cases, there seem to be no strict grammatical constraints which would uniquely determine the choice between *-(e)k* and *-le(ŋ)* for F-marking and, accordingly, between *-hane*∼*-qane* and *-le(ŋ)* for non-focal P-marking. However, there is a very strong statistical correlation between this choice and the presence of non-possessive modifiers: NPs without modifiers tend to receive *-le(ŋ)* as their F- and P-marker (see (1), (4)–(6)), and NPs with non-possessive modifiers,[4] *-(e)k* for F-marking and *-hane*∼*-qane* for P-marking (see (7), (12)).

(12) *mārqa-n t'am-uo-d'e jalhil-ek l'e-l*
one-AT big-STAT-AT lake-P|F be-SF
'There was A BIG LAKE.'

These tendencies are violated in about 5 per cent of cases, and the factors that trigger these violations are not entirely clear.

Note that the distinctions between these case paradigms do not amount to differences in phonological forms of case markers; they mark different functional oppositions. The third person pronouns, proper names, and possessive NPs allow for a single morphological opposition, the locative marker for non-focal Ps vs. the unmarked form for all other core NPs. This can lead to a violation of the basic constraint that requires P to be explicitly marked in the context of a third person A and thus ensures discrimination between A and P in transitive clauses: in a P-focus clause with P represented by an NP from this class, both core arguments remain unmarked:

(13) *ma:rquo-d'e mit uo korel bun'i-l-ŋin' l'e-mle*
one-AT our child ogre kill-ANR-DAT AUX-PF.3SG
'An ogre is going to kill OUR ONLY CHILD.'

Although the most frequent word order in transitive clauses is APV, the relative order of A and P can be easily reversed (as shown by (13)), so the listener cannot rely on it for disambiguation of argument structure. However, this clause is unambiguous, because the verb marks it as a P-focus sentence, and A is represented by a bare noun, which would have to be overtly marked if it were P. A P-focus sentence would be really ambiguous only if both core arguments belonged to the class incompatible

[4] Demonstrative determiners play no role in the choice of case marker.

978–0–19–920647–6 56-Spencer-ch55 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 794 of 796 May 16, 2008 19:1

with overt F-markers (e.g. if the ogre had been referred to by his proper name in (13)). It seems, however, that such sentences are avoided by the speakers of Yukaghir; in particular, there are no such examples in the published corpus of Tundra Yukaghir texts (Maslova 2001). One avoidance strategy is illustrated in (11) above, where a proper name is encoded as a modifier of a common noun, which can (and must) receive an overt F-marker.

The paradigms for complex and simple NPs differ in whether the opposition between F and non-focal P can be neutralized: *-(e)k* unambiguously marks its NP as F (and as P in a transitive clause), whereas the F-marking meaning of *-le(ŋ)* is neutralized in the context of third person A. It should be noted, however, that the statistical distribution between two variants of this suffix, *-leŋ* and *-le*, is sensitive to the actual function of the suffix. The following figures are based on the corpus published in Maslova (2001):

(14)

	-leŋ	*-le*
S-focus	96%	4%
P-focus	60%	40%
non-focal P	2%	98%

These figures suggest that *-leŋ* correlates with the F-marking function, and *-le*, with the P-marking function. If these functions are combined, there is no strong preference for one or the other variant.

55.4 Kolyma Yukaghir: a regularized case paradigm

This distribution shown in (14) gives a convenient starting point for a brief comparison between Tundra and Kolyma Yukaghir, where the F-marking and P-marking functions are consistently distinguished for all NPs. More specifically, the F-marking counterpart of *-le(ŋ)* in Kolyma Yukaghir is *-le-k*, and its P-marking counterpart is *-le*. Based on the regular phonetic correspondences alone, one would predict only the second variant, since the word-final *ŋ* has been lost in Kolyma Yukaghir in all other contexts. It seems plausible to hypothesize, therefore, that the final *-k* in *-le-k* can be traced back to the other F-marker *-(e)k*, which is also retained in Kolyma Yukaghir. As a result, *-(e)k* has become a general F-marker (with the exception of NPs incompatible with overt F-marking). Further, the Kolyma Yukaghir counterpart of the Tundra Yukaghir locative P-marker is *-ge-le*∼*-ke-le*, which can be safely characterized as an accusative marker: on the one hand, the

978–0–19–920647–6 56-Spencer-ch55 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 795 of 796 May 16, 2008 19:1

corresponding locative suffix has been lost in Kolyma Yukaghir;[5] on the other hand, its final component is identical with the other P-marking suffix, *-le*. The overall result is a regularized paradigm, where the word-final suffix is associated with the external function of NP (*-le* for non-focal P and *-k* for F), and the pre-final marker (*-ge-*∼*-ke-* and *-le-*, respectively) reflects the internal structure of NP:[6]

(15) 'Accusative' (non-focal P) *-(ge∼ke)-le*
'Focus (S/P)' *-(le)-k*

At this diachronic stage, only the word-final morphemes can be identified as genuine case markers.

55.5 Emergence of focus-oriented splits in case marking

To sum up the discussion so far, the Yukaghir case-marking system is shaped by the fact that nominal focus markers are employed to distinguish P from A. This is most obvious in the case of the Tundra Yukaghir *-le*(*ŋ*), yet is also true for all other overt F-markers, since they are incompatible with A. This combination of functions appears to be uncommon cross-linguistically, yet it is certainly not unique; for example, a system very similar to the Yukaghir one is attested in some varieties of Dogon (Culy 1995; Plungian 1995). Such systems can be accounted for as grammaticalized counterparts of a statistical correlation between P and F, which can be assumed to exist whenever an F-marking strategy is commonly used for so called 'broad focus' structures, where the focal part of the message is not limited to the F-marked argument itself, but can also include the verb: since A is most frequently the topic of its sentence, an in-focus argument of a transitive clause is considerably more likely to be P than A. For Yukaghir, the corresponding diachronic scenario can be briefly outlined as follows.

First, the modern P|F-markers can be traced back to the nominal predicate markers, as shown in (16):

(16) a. *taŋ tett'ie leml'e-leŋ*
that rich.man chieftain-PRED
'That rich man was a chieftain.'

[5] It is retained only in some non-finite verb forms, etymologically related to the locative of action nominal.

[6] Another salient point of divergence between the Yukaghir languages is the absence of a special A-focus construction in Kolyma Yukaghir (the A-focus information structure is expressed by (P)AV word order and intonation).

978–0–19–920647–6 56-Spencer-ch55 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 796 of 796 May 16, 2008 19:1

b. *ki-d* *akā-pe-gi* *qāli-t'e* *ierut'e-pe-k*
two-AT elder.brother-PL-3 formidable-AT hunter-PL-PRED
'His two elder brothers were formidable hunters.'

c. *met* *peldudie* *qojd-enie-gi*
my husband god-mother-3
'(She was) my husband's godmother.'

There are only two minor differences between the modern P|F-markers and the nominal predicate forms: first, the final *ŋ* in *-leŋ* is obligatory in the latter function; secondly, the dependency on the internal structure of NP is strictly deterministic for the nominal predicate markers (in contrast to the essentially statistical dependency described above for the P|F-markers). Further, the S-focus and P-focus verb forms are transparently related to the verb forms used in relativization. This suggests that the diachronic source of focus constructions is a cleft-like focus-marking strategy, with the focal NP encoded as a nominal predicate, and the remainder of the sentence as a relative clause.

Although the hypothesized cleft-like strategy might have been grammatically constrained from the very beginning because of the general constraints on relativization (which could have excluded peripheral NPs), it could hardly be incompatible with A. The emergence of this new grammatical constraint can be explained as a result of functional extension of the F-marking strategy to the domain of broad focus structures, which must have significantly increased the discourse frequency of P-focus sentences. At the present time, for example, *c.*97 per cent of all transitive sentences with nominal focus are P-focus sentences. Thus, even in the absence of grammatical constraint against overt F-marking of A, the F-markers would be linked to the P role in the overwhelming majority of transitive sentences containing such a marker, so that the listener could be almost sure that the focal NP is P. Assuming there were no other case markers in place to counteract this 'default' interpretation, the speakers would then avoid using this F-marking strategy for A in order not to mislead the listeners, thereby reinforcing the link between the F-marking and P. This path of development naturally leads to a gradual disappearance of the cleft-like A-focus construction, so that the P-marking function of F-markers is integrated into the grammar. Apparently, the F-marker of simple NPs was subsequently reanalysed as a general P-marker. The F-marker of complex NPs didn't undergo such reanalysis and cannot be used for non-focal Ps in either Yukaghir language. The use of locative for P-marking is obviously intrinsically related to this difference in diachronic paths of the F-markers, yet whether it was a cause or an effect of this difference remains an open question.

CHAPTER 56

CASE RELATIONS IN TLAPANEC, A HEAD-MARKING LANGUAGE

SØREN WICHMANN

Tlapanec is a head-marking language, in which case relations are marked by verbal suffixes. Four types of relations are marked in this way. Three of these bear many resemblances to cross-linguistically well-known cases: the Ergative, Absolutive, and the Dative. The fourth, however, is a novel type of relation for which I have had to coin a neologism: the Pegative. This encodes an actor involved in an event which also involves a Dative-like undergoer. This paper will present the Tlapanec system with special attention to the three following features: (1) the markers attach to the predicate, (2) the Ergative is morphologically unmarked, (3) the inventory includes the novel Pegative relation.[1]

[1] Data in this paper are from the Azoyú variety of Tlapanec (autodesignation: me$ʔ^{M}$pa^{H}). The language is spoken by a few hundred speakers in Guerrero, Mexico, and belongs to the Otomanguean family. I would like to acknowledge helpful comments from Andrej Malchukov and an anonymous internal referee.

978–0–19–920647–6 57-Spencer-ch56 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 798 of 807 May 16, 2008 19:1

56.1 Typological profile

By way of a brief typological profile the following characteristics may be highlighted. Tlapanec exhibits a VAO basic word order with the possibility of fronting A or O in a topicalization construction. Predicates inflect for aspect, polarity, and person, while nouns may inflect for person of possessor. An agentive–patientive distinction is expressed by the presence vs. absence of a specialized set of agentive prefixes: *ta-* (and allomorphs) in second person singular and *u-* in the plural (Wichmann 1996). There are seven tones: high (H), mid (M), low (L), and four contour tones (HM, MH, ML, LM), and all are found in both lexical and grammatical domains although the contour tones are rarer than level tones in the lexical domain. Predicates agree with their arguments in animacy, and the way they inflect depends on the following four different situations: the predicate does not take an animate participant, only an inanimate one (−A); the predicate takes either an animate argument or an animate and an inanimate one (A); two animate arguments plus/minus an inanimate one are involved (AA); three animate arguments (AAA) are involved. Examples (1a–b) illustrate how the verb form meaning 's/he is sprinkling' changes when the valency increases from one to two animate arguments (we discuss these examples in more detail towards the end of the chapter).

(1) a. *nandre*^{L} *hme*^{H} *i*^{M}*ya?*^{M} 's/he is sprinkling water'
s/he.sprinkles water
b. *nandri*^{M}*hma*^{H} *i*^{M}*ya?*^{M} 's/he is sprinkling water on him/her'
s/he.sprinkles water

An animacy hierarchy animate > inanimate determines which participant is indexed on the verb. In (1a), for instance, the indexed participant is the human agent. In addition, a role hierarchy actor > undergoer > theme is relevant. For A and AA verbs that include an inanimate argument the animate participant(s) always rank higher than the inanimate participant on the role hierarchy. Thus, an expression like 'the hammer killed the man' is impossible. The highest-ranking third person animate participant on the role hierarchy is cross-referenced on the verb for a given vs. new distinction, which bears some resemblance to both obviation and switch-reference (Wichmann, in press). This distinction is also expressed by pronouns, possessed nouns, and numerals, which are all elements that behave morphologically, although not syntactically, like predicates. There are no passive or antipassive constructions in the language, the nearest equivalents of passives being resultatives or impersonals. Thus, there are no arguments for positing grammatical relations in Tlapanec. There are no adjectives in the language, property concepts being expressed by stative verbs. Apart from a highly productive iterative derivation the synchronically identifiable derivational morphology is largely restricted to some non-productive causative prefixes.

978–0–19–920647–6 57-Spencer-ch56 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 799 of 807 May 16, 2008 19:1

56.2 The morphology of case marking

Tlapanec monopersonal (A) verbs fall into four different morphological classes identified by patterns of suffixation. Sample paradigms of each are given below. The stative verb 'to be tall' is only inflected for person, the other three verbs are additionally inflected for aspect by means of the imperfective prefix *na-*.

(2) Examples of four different Tlapanec verbal paradigms

	'to throw down' (tr)	'to be tall'	'to cover' (tr)	'to pass, cross' (intr)
1	*na-hti*^{ML} *gu*^{LM}	*ɸi*^{L} *dũʔ*^{LM}	*na-ko*^{M} *go*^{LM}	*na-no*^{M} *hgoʔ*^{L}
2	*na-ta-hti*^{L} *gu*^{L}	*ɸi*^{L} *dãʔ*^{LM}	*na-ta-ko*^{M} *ga*^{LM}	*na-(ta-)no*^{M} *hgaʔ*^{ML}
3N	*na-hti*^{L} *gu*^{L}	*ɸi*^{L} *daʔ*^{HM}	*na-ko*^{M} *go*^{HM}	*na-no*^{M} *hgo*^{ML}
4i	*nu-hti*^{ML} *gu*^{LM}*=luʔ*^{M}	*ɸi*^{L} *dãʔ*^{LM}*=luʔ*^{M}	*nu-ko*^{M} *ga*^{LM}*=luʔ*^{M}	*na(/u)-no*^{M} *hgaʔ*^{L}*=luʔ*^{M}
4x	*nu-hti*^{ML} *gu*^{LM}*=loʔ*^{L}	*ɸi*^{L} *dãʔ*^{LM}*=loʔ*^{L}	*nu-ko*^{M} *ga*^{LM}*=loʔ*^{L}	*na(/u)-no*^{M} *hgaʔ*^{L}*=loʔ*^{L}
5	*nu-hti*^{ML} *gu*^{LM}*=laʔ*^{L}	*ɸi*^{L} *dãʔ*^{LM}*=laʔ*^{L}	*nu-ko*^{M} *ga*^{LM}*=laʔ*^{L}	*na(/u)-no*^{M} *hgaʔ*^{L}*=laʔ*^{L}
6N	*nu-hti*^{M} *gu*^{L}	*ɸi*^{L} *dīʔ*^{HM}	*nu-ko*^{M} *ga*^{ML}	*na(/u)-no*^{M} *hgũ*^{LM}

1: first person, 2: second person, 3N: new third person, 4i: first person inclusive, 4x: first person exclusive, 5: second person plural, 6N: new third person plural.

The different classes are distinguished by suffixes fusing the marking of person and case relations. In (3) I have extracted the segmental parts of the suffixes in question. These suffixes carry tones which interact with underlying tones of the stems to which the suffixes attach such that several tonal subparadigms for each of the four main paradigms are produced. Although some progress has been made towards understanding the nature of internal tonal sandhi (Wichmann, forthcoming) it has not been established what the underlying tones associated with the various suffixes are. To do this, one must first establish what the underlying stem tones are, but this is difficult because verbal stems are always inflected. Moreover, there are many instances of irregularity. Verbs subcategorizing for the Ergative relation exhibit several dozen different tonal paradigms, whereas verbs in the other three categories each exhibit five major tonal subparadigms.

(3) Case markers of monopersonal verbs[2]

	Ergative	Absolutive	Pegative	Dative
1	*-Ø*	*-ũʔ*	*-u* ∼ *-o*	*-uʔ* ∼ *-oʔ*
2	*-Ø*	*-ĩʔ/-ãʔ*	*-a/-i*	*-aʔ*
3N	*-Ø*	*-i/-a*	*-u* ∼ *-o*	*-u* ∼ *-o*
4–5	*-Ø*	*-ãʔ*	*-a/-i*	*-aʔ*
6N	*-Ø*	*-ĩ*	*-a/-i*	*-ũ*

[2] 'N' stands for 'new' and contrasts with 'G' for 'given' and is, as mentioned in section 56.1, similar in many respects to obviation (and to a lesser degree to switch-reference). Although morphologically the New form is more basic, I normally use as citation form the third person singular Given, the

Many scholars (e.g. Blake 2001) either overtly state or implicitly assume that case marking is a phenomenon restricted to nouns, whereas others (e.g. Butt 2006: 5–6) do not insist on such a requirement. Finally, a few (e.g. Kibrik, in press) explicitly expound the view that pronominal markers in head-marking languages (e.g. Navajo) do carry case marking. According to the latter view, there is no a priori reason to assume that the category of case could not be expressed on either the head, the dependent, or both in different languages, since languages preferentially either mark relations between a head and a dependent within a phrase on the head, on the dependent, or on both (Nichols 1986). Blake's definition, according to which case is a category which 'marks the relationship of a noun to a verb at the clause level' (Blake 2001: 1), requires a noun to be present in the clause. In a situation of head-marking, where the referential status of pronominal markers (Jelinek 1984, Mithun 1985) obliterates the need for overt noun phrases except when new participants are introduced, case does not necessarily mark the relationship of a noun to verb. Rather, what is more generally true of case marking is that it signals the relationship of an argument to a predicate at the clause level. Adherents of the opposite view that case marking is a phenomenon restricted to nouns customarily describe potential candidates for case marking pronominal affixes attached to verbs under the rubric of agreement. Corbett (2003: 166) states that 'whether incorporated pronouns are near the edge of the phenomenon of agreement or the beginning of the neighbouring phenomenon is less important than seeing the connections and differences'. In a similar vein, I will approach the Tlapanec verbal suffixes with a view to their similarities and differences with case markers and leave it to the reader to decide on the rather uninteresting terminological question of whether the phenomenon should or should not be called 'case' proper.

One difference with respect to canonical case markers in dependent-marking languages is that the Tlapanec verbal suffixes have a portmanteau behaviour and also mark person distinctions. Contributions to person marking are made by other parts of the grammar as well, however. For agentive verbs, a contribution comes from the second person agentive prefix *ta-* or allomorphs thereof – among the examples in (2) 'to be tall' is patientive, 'to throw down' and 'to cover' are agentive, and 'to pass across' is 'fluid' in the sense that it may be treated as either agentive or patientive. Another contribution comes from the enclitics $=lu\text{ʔ}^{M}$, $=lo\text{ʔ}^{L}$, and $=la\text{ʔ}^{L}$, which specialize in distinguishing among plural speech act participants. For bi- or tri-personal transitive (AA or AAA) verbs there is not a one-to-one relationship between the presence of a given suffix from the paradigm in (3) and a tone pattern, which suggests that tones are added to mark agents and that these added

reason being that this form can constitute a whole sentence in itself, whereas the N form must be followed by an overt mentioning of the pivot argument of the verb. The G form is derived from the N form by tonal affixation and, for verbs subcategorizing for the Ergative, additionally by a suffix *-i* which merges with the stem vowel. The symbol '∼' indicates predictable morphophonological alternation, the symbol '/' an alternation which is to a lesser degree predictable, although it is clearly also phonologically conditioned, at least historically.

tones enter into interaction with the tones resulting from the sandhi processes involving stem tones and tones associated with the suffixes. For lack of space these complex paradigms cannot be illustrated and discussed here.

The Tlapanec markers are also unusual in comparison to canonical case systems inasmuch as the Ergative is the unmarked member of the set. It is a well-known fact that markedness relations in case systems are generally such that the morphologically and functionally unmarked member tends to be the nominative in accusative languages and the absolutive in ergative languages. This observation was first made by Greenberg (1963: 75) and has been elaborated on by Dixon (1994: 63–96) and others. Nevertheless, 'marked nominative' systems are found in many African languages, particularly in some subgroups of the Afroasiatic family, including most Cushitic languages, most Omotic languages, and several Berber languages or dialects. It is also found in some languages of the Nilotic subgroup of Nilo-Saharan, which is geographically contiguous to Cushitic. 'Marked nominative' also occurs in some languages of California and adjacent regions, as well as sporadically elsewhere in the world (see König, Chapter 35). There are far fewer exceptions to Greenberg's generalization for languages that have ergative–absolutive alignment. One case of a language reported to exhibit 'marked absolutive' is the Austronesian language Nias (Donohue and Brown 1999, Brown 2001). Within the Otomanguean language family, to which Tlapanec belongs, the phenomenon has been noted for Sochiapan Chinantec (Foris 2000: 7, 254). Thus, the marked absolutive of Tlapanec is unusual when interpreted as a case marker, but it does not violate any absolute language universal.

We may now turn to the functional similarity between the Tlapanec markers and more canonical case marking systems. As we deal with the morphosyntax of the Tlapanec markers and their semantics in the following two sections, we shall see that their function is precisely to signal the relationship of an argument to a predicate at the clause level.

56.3 The mechanisms of case assignment

Verbs assigning the Absolutive are mostly intransitive, including stative verbs. Verbs assigning the Ergative and Pegative are generally transitive, and verbs assigning the Dative may be either transitive or intransitive. Furthermore, the Dative is used for marking nominal possession, as demonstrated in the following paradigms:

(4)					
	biʔ^{M}i^{M} 'day'	*i^{M}štuʔ^{M}* 'basket'	*mi^{M}ša^{M}*	'shadow'
1	*biʔ^{M}yuʔ^{L}*	*i^{M}štuʔ^{L}*	*mi^{M}šoʔ^{L}*	'my day / basket / shadow'
2	*biʔ^{M}yaʔ^{ML}*	*i^{M}štaʔ^{ML}*	*mi^{M}šaʔ^{ML}*	'your . . . etc.'
3G	*biʔ^{M}yu^{H}*	*i^{M}štuʔ^{H}*	*mi^{M}šo^{H}*	

3N	*biʔ^{M}yuʔ^{ML}*	*i^{M}štu^{ML}*	*mi^{M}šo^{ML}*
4/5	*biʔ^{M}yaʔ^{L}=*	*i^{M}štaʔ^{L}=*	*mi^{M}šaʔ^{L}=* (add *=luʔ^{M} =loʔ^{L},=laʔ^{L}*)
6G	*biʔ^{M}yũ^{MH}*	*i^{M}štũʔ^{MH}*	*mi^{M}šũ^{MH}*
6N	*biʔ^{M}yũ^{LM}*	*i^{M}štũʔ^{LM}*	*mi^{M}šũ^{LM}*

The use of the Dative for marking possession is not uncommon cross-linguistically.

When there are two animate arguments one will be the agent and the other the patient or one will be the source/stimulus and the other the recipient. (I use terms for semantic roles in a broad sense that does not imply strict conformity to semantic criteria of assignment). Corresponding to these two fundamental types of relations there are two types of paradigms for dipersonal verbs (i.e. verbs taking two animate arguments),[3] which are shown in (5) below. Verbs involving the agent–patient relation are organized in an ergative pattern since the Absolutive endings refer to the patient. For the source/stimulus–recipient relation the organization is split ergative in the person dimension. When a non-third person singular recipient is involved, the verbs take the Dative endings, referring to the recipient (the one

(5) Case markers of monopersonal and dipersonal verbs

Monopersonal verbs				Dipersonal verbs		
Person	Absolutive	Dative	Pegative	Person combination	Absolutive	Dative/ Pegative
1	*-ũʔ*	*-uʔ/-oʔ*	*-u/-o*	2/3N/5/6N-**1**	*-ũʔ*	*-uʔ/-oʔ*
				3G/6G-**1**		*-eʔ*
2	*-ãʔ/-ĩʔ*	*-aʔ*	*-a/-i*	1/3/4x/6-**2**	*-ãʔ/-ĩʔ*	*-aʔ*
3	*-a/-i*	*-u/-o*	*-u/-o*	**1**-3 **2**-3 **3**-3 **4**/ **5**-3 **6**-3	*-a/-i*	*-u/-o* *-a/-i* *-u/-o* *-a/-i* *-a/-i*
4-5	*-ãʔ*	*-aʔ*	*-a/-i*	3/6-**4**/ **5**, 2/ 5-**4**x, 1/4x-**5**	*-ãʔ*	*-aʔ*
6	*-ĩ*	*-ũ*	*-a/-i*	1/2/3/4/5/6-**6**	*-õ*	*-u*

[3] Tripersonal (AAA) verbs are derived from the dipersonals by adding a suffix *-i* indicating the presence in the argument structure of an animate theme.

exception to this is the marker *-eʔ* for third person Given acting on first person). When a third person singular recipient is involved, the verbs take the Pegative case endings referring to the source/stimulus. The Ergative marker is never involved in dipersonal verbs.

To summarize, the inflection of dipersonal verbs expressing the agent–patient relation is organized in an ergative fashion, whereas the inflection of dipersonal verbs expressing the source–recipient relation is organized in a split ergative fashion in the person dimension such that a third person singular recipient triggers a nominative–accusative pattern, whereas combinations not involving a third person singular recipient trigger an ergative pattern.

56.4 The semantics of case assignment

Here follow a few examples of monopersonal verbs that assign each of the four case relations. When studying these examples the reader should keep in mind that the relations encoded always pertain to the animate argument.

(6) Some verbs subcategorizing for the four different case relations

ERGATIVE	*na-*^{M}*ʔdi*^{H} 's/he is sowing it'
	na-^{M}*hwe*^{H} 's/he is selling it'
	na-kaʔ^{M}*wi*^{H} 's/he is hiding it'
	na-ɸe^{M}*ke*^{H} 's/he is smoking/burning it'
	na-šuʔ^{M}*mbiʔ*^{H} 's/he is roasting something'
ABSOLUTIVE	*ba*^{L}*wi*^{H} 's/he is alone'
	da^{M}*ska*^{H} 's/he smells bad'
	hka^{M}*ma*^{H} 's/he is hung up'
	na-mya^{M}*hwĩ*^{M} 's/he is worrying'
	na-wa^{M}*pa*^{H} 's/he has time'
PEGATIVE	*geʔ*^{M}*do*^{H} 's/he has (something)'
	na-ka^{L}*šu*^{H} 's/he is skinning it'
	na-mi^{M}*nduʔ*^{H} 's/he is seeing it'
	na-ɸiʔ^{M}*yuʔ*^{H} 's/he is putting it out' (e.g. light)
	na-re^{M}*ko*^{H} 's/he is blocking it' (e.g. road)
DATIVE	*ba*^{M}*šo*^{H} 's/he is nude'
	na-mbiʔ^{M}*yu*^{H} 's/he is called (something)'
	na-^{M}*ndo*^{H} 's/he wants it'
	na-^{M}*hmyuʔ*^{H} 's/he is using it'
	na-ka^{M}*nu*^{H} 's/he is given it'

978–0–19–920647–6 57-Spencer-ch56 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 804 of 807 May 16, 2008 19:1

Ergative and Pegative, on the one hand, and Absolutive and Dative, on the other, make distinctions regarding agency, Ergative or Pegative being assigned to actors and Absolutive or Dative to undergoers. (I have based the term 'Pegative' on the Greek πηγή, which means 'origin, source, emanation, etc.', to provide a name for a case relation that prototypically refers to a giver as opposed to a recipient). The main semantic parameter that is involved in distinguishing the two different kinds of actor and the two different kinds of undergoer seems to be one of the degree of impact of the action, that is, an effectedness/affectedness parameter. For verbs assigning the Pegative the effect generally seems to be lower than for verbs assigning the Ergative. Often the undergoer is only partially affected. Thus, 'to sow', 'to sell', 'to hide', 'to smoke', 'to roast something' have a direct impact and/or involve the undergoer as a whole, whereas 'to have', 'to skin', 'to see', 'to put out', 'to block something' imply a lesser or partial effect. The verbs assigning the Absolutive often describe more permanent states than verbs assigning the Dative. Thus there is a mirror relationship where Ergative–Absolutive are each other's opposites just like Pegative–Dative are each other's opposites. The relationships are summarized in (7).

(7) Semantics correlates of Tlapanec case assignment

MACRO-ROLE / E(/A)FFECTEDNESS	ACTOR	UNDERGOER
HIGH	ERGATIVE	ABSOLUTIVE
LOW	PEGATIVE	DATIVE

Verbs that take Ergative arguments are nearly all transitive (see examples in (6)), although a few intransitives have been recorded, including *-mbe^{M}ye?^{H}* 'to cry', *-ra?^{M}tsi?^{H}* 'to sneeze', *-ru^{L} mba?^{MH}* 'to work as a day labourer', *^{M}-?sŋe^{H}* 'to dance', *-ŋgo^{M}wẽ?^{H}* 'to scream', *-ŋgu^{L} nda^{H}* 'to dream'. Verbs subcategorizing for the Absolutive comprise both dynamic and stative verbs, that is, verbs that respectively do and do not inflect for aspect. The dynamic verbs include a handful of transitives taking inanimate patients, e.g. *-hka^{H}ma^{H}* 'to hang up', *-hya^{L}wĩ?^{M}* 'to scratch', *-nde^{M}mbo^{MH}* 'to gather', *-ra^{M}šno^{H}* 'to count', *^{L}-?ye?^{H}* 'to look for', but the great majority are intransitives (examples were already given in (6) above). All verbs that I have recorded which take the Pegative are transitive. About two-thirds of the verbs taking Dative participants are intransitive (be they dynamic or stative) and a third are transitives. Some examples of the minority category were already given in (6), and some others follow here: *-ga^{M}hŋgo?^{H}* 'to swallow', *-kra?^{M}ŋgo^{H}* 'to

carry something crosswise', *-ma*^{H}*yũ*^{LM} 'to be able to, learn', ^{M}*-mbo*^{H} 'to finish, use up', *-mbo*^{M}*-mo*^{H} 'to forget', *-ga*^{L} *hŋgũʔ*^{MH}. The general lack of a match between the marking of case relations and transitivity is not surprising given that 'transitivity', in the sense of how many objects a verb may take, is not central to the organization of Tlapanec grammar. As mentioned, what is crucial is rather the number of animate participants allowed by the valency of a given verb.

Obviously the semantic characterizations summarized in (7) must remain approximations. Apart from the addition to the case inventory, the Pegative, the system does not diverge functionally to any great degree from commonly attested case systems. When we include dipersonal verbs in the discussion, the Tlapanec system begins to show its dynamicity and yet more parallels with canonical case systems turn up.

Dipersonal verbs take either Absolutive or Dative/Pegative (the latter being subject to a split pattern in the person dimension, as explained above). In (8) I provide examples of Absolutive-assigning and Dative-assigning dipersonal verbs. The undergoer-participant cross-referenced on the verb is always the indirectly or partially affected animate argument. Although one should remember that Tlapanec does not have grammatical relations, the system may essentially be equated with primary object languages (Dryer 1986), which rank indirect objects higher than direct objects for purposes of indexing.

(8) Examples of case-assigment of dipersonal verbs

ABSOLUTIVE	*-kugraʔ*^{M}*a*^{H} 'to lock up someone'
	-čiˑ^{L} *hpa*^{H} 'to hug someone'
	-guhpra^{L} *ʔa*^{H} 'to kick someone'
	-hmara^{M}*wīʔ*^{H} 'to greet someone (by caressing)'
	-hmiˑ^{L} *da*^{H} 'to shoot someone'
	-hŋga^{M}*wi*^{H} 'to protect, take care of someone'
	-kaʔ^{M}*wi*^{H} 'to hide someone'
	-ɬi^{M}*-hi*^{H} 'to make someone stand up'
	-šiˑ^{L} *ya*^{H} 'to kill someone'
	-škuʔ^{M}*-ŋgaʔ*^{H} 'to push someone'
DATIVE	^{L}*-šnu*^{H} 'to give (something) to someone'
	-ni^{M}*yũʔ*^{H} 'to leave someone'
	-nja^{M}*ũ*^{H} 'to listen to, obey someone'
	-raʔ^{M}*nu*^{H} 'to meet someone'
	^{L}*-sko*^{H} 'to chase away someone'
	^{M}*-ʔsŋgo*^{H} 'to teach someone something'
	^{M}*-ʔtũ*^{H} 'to tell someone something'
	-šna-teʔ^{H}*yõ*^{H} 'to borrow (something) from someone'
	-šneʔ^{H}*do*^{H} 'to load (e.g. an animal) (with something)'
	^{M}*-ʔyo*^{H} 'to see someone'

There are several examples where one and the same verb may assign different case relations. One example is the verb 'to sprinkle' which includes among its instantiations[4] the three forms illlustrated in (9). The examples in (9a) and (9c) are repeated from (1a, b), where they were introduced with limited glossing. At this point the reader is better equipped to appreciate detailed glosses.

(9) a. [Monopersonal, Ergative]
na-ndre^{L} *hme*^{H} *i*^{M}*ya*ʔ^{M}
IPFV-sprinkle.3G.ERG[5] water
'S/he is sprinkling water.'

b. [Monopersonal, Pegative]
na-ndri^{L} *hm-u*^{H} *i*^{M}*ya*ʔ^{M} *i*^{M}*n-u*^{ML} *ša*^{M}*bu*^{L}
IPFV-sprinkle-3G.PEG water face-3N.DAT[6] man
'S/he is sprinkling water on the face of the man.'

c. [Dipersonal, Absolutive]
na-ndri^{M}*hm-a*^{H} *i*^{M}*ya*ʔ^{M}
IPFV-sprinkle-3G>3ABS water
'S/he is sprinkling water on her/him.'

The monopersonal, Ergative-assigning instantiation *-ndre*^{L} *hme*^{H} means 'to sprinkle something' (9a). The Ergative case implies an Absolutive-like undergoer. In contrast, the monopersonal, Pegative-assigning instantiation *-ndri*^{L} *hmu*^{H} means 'to sprinkle something onto something' (9b). It acquires this meaning not because there is some valency-augmenting mechanism and/or external possession involved but because the Pegative actor implies a Dative-like undergoer, that is, a partially affected undergoer. In (9c) we see a dipersonal instantiation.

There are several examples of verbs whose instantiations respectively assign Absolutive and Ergative case. The difference is one of 'transitivity', but again not induced by a valency-changing process but rather by the semantics of the cases, which dictate that a monopersonal verb assigning the Absolutive is normally intransitive and a monopersonal verb assigning the Ergative is normally transitive. Some pairs are shown in (10).

[4] I use 'instantiation' because it is not clear that there is an underived base form from which other forms are derived. Instead, verbs may be seen as being based on general, abstract templates of which the various inflectional forms are instantiations.

[5] *ndre*^{L} *hme*^{H} consists of a stem of segmental shape *ndrihma* which has unknown underlying tones and is inflected for person by tones and for the Given category by a suffix *-i*, which merges with the final stem vowel, producing an *-e* which subsequently triggers harmony with the preceding vowel.

[6] Dative-marking here does not signal a relation between 'face' and the predicate 'to sprinkle', but rather the relation between the possessed item, 'face', and the possessor, 'man'. The verbal argument is the whole noun phrase 'the man's face'.

978–0–19–920647–6 57-Spencer-ch56 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 807 of 807 May 16, 2008 19:1

(10) Examples of pairs of instantiations of A verbs that assign Absolutive vs. Ergative

a. *-hpaʔ^{L} ã^{M}* 'to stick one's head out' [Monopersonal, Absolutive]
-hpaʔẽ$^{L\,H}$ 'to throw something inside' [Monopersonal, Ergative]

b. *-hmi^{L} di^{H}* 'to burst' [Monopersonal, Absolutive]
-hme^{L} de^{H} 'to make something burst' [Monopersonal, Ergative]

c. *-hpraʔ^{M}a^{H}* 'to enter quickly' [Monopersonal, Absolutive]
-hpraʔ^{L}e^{M} 'to chop something in two' [Monopersonal, Ergative]

Pairs of instantiations taking respectively an Absolutive and a Dative human undergoer should exist. For instance, we would expect to find a difference between, say, 'to instruct someone' (Absolutive) as opposed to 'to teach someone something' (Dative). Possibly due to limitations in my data, which are all drawn from texts (except paradigms which obviously must be elicited directly) I have not yet recorded clear instances of this.

56.5 Conclusion

In the above I have argued that a set of Azoyú Tlapanec verbal suffixes bear strong similarities to canonical case marking. I tried to demonstrate that the assignment of case relations operates on a semantic basis quite similar to what is standardly expected from a case system even if the Tlapanec system has some structural peculiarities – including a type of case relation apparently not attested in other languages, namely the one for which I have coined the term Pegative. One of the striking parallels to cross-linguistically common case systems is that the Dative is involved in marking possession. Moreover, I showed that one and the same verb may assign different relations, something which indicates that the system indeed serves to distinguish different relations between predicates and arguments at the clause level, as do case-marking systems universally. The typologically unusual morphologically unmarked status of the Ergative was briefly discussed. As it turns out, this feature is not only attested elsewhere in the world, but even recurs within the Otomanguean family of languages to which Tlapanec belongs.

CHAPTER 57

'CASE RELATIONS' IN LAO, A RADICALLY ISOLATING LANGUAGE

N. J. ENFIELD

57.1 Introduction

By definition, a language of the isolating type (Sapir 1921: 126) entirely lacks explicit morphological marking of grammatical relations or other formal linking of predicates and arguments, whether this marking be on clausal heads (agreement) or dependents (case-marking). This chapter looks at data from Lao, a radically isolating Southwestern Tai language spoken in Laos, Thailand, and Cambodia, and asks how speakers of such a language might cope without case. Does it find alternative means to the same functional ends? Does it simply leave these problems unsolved? Two points are made here from the Lao data. First, the informational problem of disambiguating role and reference of arguments hardly needs a formal solution – that is, there need be no formal alternative to case-marking, in languages which lack case. Where case-marking simply distinguishes who from whom, it is

mostly dispensable, thanks to the richness of pragmatics. A second finding is that for more 'expressive' functions of case-marking, where features of transitivity are manipulated for expressive or information-structural effect,[1] Lao finds constructional means to treat certain arguments in special ways, thereby explicitly marking non-redundant semantic information in case-like ways.

57.2 Pragmatically oriented grammar: Context as the main referential guide

As defined by its isolating morphological profile, Lao lacks both case and its functional cousin, agreement. Lao is like Mandarin, Thai, Vietnamese, and Riau Indonesian in exemplifying the extreme of pragmatically oriented grammar (cf. Gil 2005a). With argument–predicate relations marked neither on heads nor dependents, how to tell who from whom? A widely presumed answer is that hearers of languages like Lao are forced to rely on strict constituent order to maintain informational coherence in predicate–argument relations. This claim is, however, weak at best, since extensive ellipsis and movement create widespread surface ambiguity, without compromising communication. Accordingly, for Mandarin, Li and Thompson (1981: 26) state that 'no basic word order can be established'. On Riau Indonesian, Gil (2005b) shows that there are 'no distinctions between major syntactic categories'. He argues that observed word order patterns in the language are 'epiphenomenal' (cf. also LaPolla 1993). With this in mind, consider the Lao A/S-V-O constituent order pattern, perhaps the closest to a 'pragmatically unmarked' pattern:[2]

[1] I distinguish here between 'disambiguating' versus 'expressive' functions of core case-marking. Disambiguating functions serve the resolution of referential ambiguities in communication, most importantly helping hearers to track protagonists through discourse, and to map event/discourse participants onto distinct semantic roles or grammatical relations. Disambiguating functions distinguish who from whom, and little more. Expressive functions of core case, by contrast, signal distinctions in conceptual representation or construal of events, marking constructions which may signal special distinctions in aspect, agentivity, responsibility, involvement, and effect. An example is from Finnish, where partitive instead of accusative case-marking on an undergoer signals that the undergoer is 'only partially affected by the action' (Comrie 1989: 127). Expressive case functions tend to be optional, as distinct from the typically obligatory nature of disambiguating case-marking.

[2] By 'unmarked' I mean that speakers may report an impression that the S/A-V-O pattern is somehow basic in status. A consultant will likely supply the S/A-V-O pattern when asked to compose sentences of the variety *The farmer kills the duckling*, i.e. decontextualized structures of the sort which seldom actually occur (Du Bois 1987). The impression of basicness to this word order arises not from an asemantic structural default, but from the normal discourse asymmetry inherent in argument structure (e.g. one argument will, all things being equal, be higher on the scale of animacy, agency,

978–0–19–920647–6 58-Spencer-ch57 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 810 of 820 May 16, 2008 19:1

(1) NP_A V NP_O
kuu3 jaan4 mùng2
1SG.B afraid 2SG.B
'I was afraid of you.'

(2) NP_A V NP_O
phuø-pên3-mia2 khòòng3 thaaw4 nan4 hên3 qavaj2ñavaq1
person-be-wife of young.man DEM.NPROX see organ
khòòng3 faaj1 coon3
of side bandit
'That young man's wife saw the bandit's genitals.'

(3) NP_S V
saam3 khon2 taaj3
three person die
'Three people died.'

Departures from the A/S-V-O pattern are common. 'Movement', for example, may see a subject argument in a post-final position (examples 4 and 5, where the back slash represents the onset of a prosodic mark-off, with lowered amplitude and pitch), or an object argument in initial position (example 6, where the forward slash represents the syntactic border between a left-positioned topic and a grammatical subject). It is important to note that while it is formally apparent that something has been 'moved', there is no information about the semantic/functional role of the moved argument.

(4) V NP_S *taaj3 lèèw4 \ phòò1 hanø*
die PRF father TOP.DIST
'(He)'d be dead, the father.'

(5) V NP_O NP_A *qaw3 mia2 \ haw2 niø*
take wife 1.FAC TOP
'Took a wife, I (did).'

(6) NP_O NP_A V *lot1 / haw2 laø bòø mii2*
vehicle 1.FAC PCL NEG have
'A car, I didn't have.'

Another reason surface strings might not show canonical constituent order patterns is ellipsis. Arguments may be freely ellipsed in Lao when contextually retrievable (or 'definite'), leaving literally zero material for the mapping of arguments onto predicates, and no coherent way of determining 'constituent order'. This is perhaps the most important challenge to a claim that without case or agreement, word order is crucial for maintaining role and reference relations:

topicality, than the other; cf. Comrie 1989: 127 on this as 'natural information flow'; cf. Hopper and Thompson 1980, Langacker 1987, LaPolla 1993, Croft 2003).

(7) *ñaaw2*
long
'(It was) long.'

(8) *lùùm2*
forget
'(I have) forgotten (it).'

(9) *hên3*
see
'(I) saw (it).'

The referential resolution of ellipsis in Lao is in general completely open to pragmatic interpretation since there are few strict grammatical constraints on the interpretation of unexpressed nominal material. Consider the following example of 'gapping' (constructed, modelled after a Thai example in Foley and Van Valin 1984: 194):[3]

(10) *tam3* *khuaj2* *taaj3*
crash.into buffalo die
i. '(S/he) crashed into a buffalo and died.'
ii. '(S/he) crashed into a buffalo and (it) died.'
iii. '(S/he) crashed into a buffalo and (the car) died (i.e. stalled).'

When we combine ellipsis with movement, further structural ambiguity arises (as pointed out for Mandarin Chinese by Chao 1968; cf. Gil 2005a):[4]

(11) Surface sequence: **NP $V_{bivalent}$**
Structural analysis 1: **NP_O $V_{bivalent}$** (NP_A ellipsed)
Structural analysis 2: **NP_A $V_{bivalent}$** (NP_O ellipsed)
e.g. with bivalent verb *qaw3* 'to take':
phuak4 *juu1* *nam2* *thaang2* *kaø* *qaw3*
group be.at accompany road TLNK take
i. 'Those$_i$ along the road, (they$_j$) took ø$_i$.' (actual reading in original context)
ii. 'Those along the road took (them/it).' (possible reading)

[3] The only case of strict referential control of a 'zero' element is in same-subject readings of 'want' complements (e.g. *man2 jaak5 khaa5* [3sg want kill] 'S/he wants to kill (it/them)'; cf. *man2 fan3 vaa1 khaa5* [3sg dream COMP kill] 'S/he$_i$ dreamt s/he/it/they$_{i/j}$ killed her/him/it/them$_{j/i}$.'). Because of this syntactic control constraint in 'want' complements, it is necessary to explicitly mark switch-reference in such constructions with a dummy causative in *haj5* 'give':

(a) *man2* *jaak2* *paj3*
3SG.B want go
'S/he$_i$ wants Ø$_{i/*j}$ to go.'

(b) *man2* *jaak2* *haj5* *paj3*
3SG.B want give/cause go
'S/he$_i$ wants ø$_{*i/j}$ to go.' (i.e. 'S/he wants him/her/them to go.')

[4] For clarity of presentation, I do not include in the example structure's schematic representation the various particles which appear in the actual examples. In neither case does the presence of the particle bear upon the mapping of arguments to one or another semantic, grammatical, or discourse function.

(12) Surface sequence: **NP $V_{bivalent}$ NP**
Structural analysis 1: **NP_A $V_{bivalent}$ NP_O** (NP_O postposed)[5]
Structural analysis 2: **NP_O $V_{bivalent}$ NP_A** (NP_A postposed)
e.g. with bivalent verb *mak1* 'to like':
tamluat5 / mak1 dêj2 \ phuø-saaw3 tòòn3 nan4
police like FAC CT.HUM-girl time DEM.NPROX
i. 'Police$_i$, (they$_j$) liked (them$_i$) you know, girls$_j$ back then.' (actual reading in context) [i.e. girls liked police.]
ii. 'Police liked (them) you know, girls back then.' [i.e. police liked girls.]

These everyday Lao examples show variable or indeterminate constituent order. Such patterns are readily analysed as arising from movement and ellipsis, though note that these are merely descriptive: nothing in the form of these examples serves to disambiguate. That these ubiquitous relaxations of the 'word order patterns' co-exist with a total lack of morphological marking of semantic roles or grammatical relations might suggest chaos. But in real contexts of usage, Lao speakers have no difficulty in communicating.

The conclusion is that Lao and other radically isolating languages (Gil 2005a, Enfield 2005) demonstrate that the merely disambiguating functions of case are so redundant as to be almost entirely dispensable. When core referential information is not symbolically encoded in grammar, potential ambiguities in role/reference relations are readily resolved by features of context. Within 'context' here we may include, on the one hand, selectional restrictions of verb/argument semantics (e.g. if I give you a predicate 'eat<x_{eater},y_{eaten}>' and two arguments 'John' and 'an apple', chances are you will correctly guess the argument–role mapping), and, on the other hand, the pragmatic constraints of expectation supplied by any given active discourse world (and following informational principles of topic continuity, etc.). With these semantic and pragmatic devices alone, speakers of Lao can readily distinguish who from whom in the absence of the kind of unambiguous marking which morphological case might otherwise provide. The surface ambiguities just outlined are normally unproblematic when there is a full discourse context. Importantly, however, the same surface strings which are easily understood in situ may, if taken out of context, be impossible to interpret with referential certainty.[6]

[5] We are justified in saying that the NP is postposed, since it appears after the sentence-final particle *dêj2*, and thus in a prosodically marked-off right-position to the clause. But its postpositioned status has no bearing upon a hearer's interpretation of its semantic role or grammatical relation.

[6] A good example is (12), above. I presented an audio recording of this utterance alone, out of context, to a number of Lao speakers, and asked them to explain their interpretation of the sentence. In all cases, they understood the utterance to have the opposite mapping of actor and undergoer to what was intended by the speaker in the original context – i.e. all hearers assumed the initial noun phrase to be the A.

57.3 Argument–predicate relations: some patterns

I have stressed the lack of dependable formal marking of predicate–argument relations in Lao, but this does not imply a lack of systematicity in the mapping of arguments to predicate roles/functions. There are underlying classes of argument structure pattern which constrain the possibilities. The key patterns turn on distinctions in the semantics of verbs – that is, patterns of argument (a)symmetry and aspectual structure inherent to the semantics of distinct (classes of) verbs.

57.3.1 Monovalent predicates

For monovalent predicates, disambiguating who from whom is not an issue. There are, however, a range of different possible conceptual/semantic mappings of the relation of argument to verb for monovalent (single-argument) predicates, and hearers must be able to determine which from among a number of possible roles an argument may have. Here are some basic types of relation between a predicate and its single argument in Lao:

(13) *Active monovalent relation* (single argument S = agent/theme)
Meaning: 'S does V'; includes typical active monovalents (e.g. *caam3* 'sneeze', *lèèn1* 'run', *san1* 'shake').

(14) *Inchoative-stative monovalent relation* (single argument S = theme)
Expresses the meaning 'S enters into and/or is in state V'; typical stative 'property concept' monovalents (e.g. *laaj2* 'striped', *hòòn4* 'hot', *dii3* 'good'); inchoative reading is encouraged by irrealis or progressive marking.

(15) *Resultant state monovalent relation* (single argument S = patient/theme)
Telic agent-controlled verbs with patient/theme as subject and where agent is unexpressed and indefinite/non-retrievable (e.g. *kaang3* 'to be hoisted', *pia3* 'to be platted', *tom4* 'to be boiled').

Verbs encoding these three types of argument–verb relation differ in grammatical behaviour, notably in terms of permissible alternations. For example, for (13), no transitive alternation is possible; for (14), a caused state alternation is possible (see below); for (15), negation requires *dajø*; a transitive alternation is possible (see below). (For full details on these alternations, see Enfield in press, a.) Of more direct relevance to case-marking as a device for dealing with referential ambiguities are patterns of relation between *two* arguments and a single predicate. We turn now to those.

57.3.2 Symmetric and other non-oriented bivalent predicates

When a bivalent predicate is symmetric, in the sense that its two arguments are involved in the event in the same way and to the same degree, there is (perhaps trivially) the possibility to allow any argument in any position, without (truth-conditionally) affecting predicate–argument mapping. Take verbs of likeness: *John resembles Bill* entails *Bill resembles John* (while *John scratches Bill* does not entail *Bill scratches John*). There are also asymmetric predicates like *khaat5* 'lacking (sth.)' and *têm3* 'full (of sth.)' which show similar variability. This is because, despite being asymmetric, they are 'non-oriented', that is they do not show an obvious DIRECTION of figure/ground asymmetry (as distinct, say, from the inherent orientation of more prototypical active verbs such as 'hit'; Langacker 1987: 209ff., Comrie 1989, Croft 1991: 184ff.). Diller (1997) has pointed this out for Thai, and the same kinds of examples work in Lao too (reminiscent of celebrated English 'case alternations' like *the garden is swarming with bees* versus *Bees are swarming in the garden*; Levin 1993, Levin and Rappaport Hovav 2005, among many others). Importantly, the Thai/Lao cases are distinct from 'swarm' type examples because the alternations do not involve any kind of morphological marking, as non-core or otherwise, on arguments. With non-oriented predicates of this kind, a single truth-conditional situation may equally well be describable by expressions of opposite constituent ordering (16a, b; 17a, b), or a single expression may have two very different truth-conditional interpretations (as in (18)):

(16) (a) *còòk5 têm3 law5*
cup full liquor
'The cup is filled (with) liquor.'
(b) *law5 têm3 còòk5*
liquor full cup
'Liquor fills the cup.'

(17) (a) *sùa4 nii4 tit3 namø-mùk2*
shirt DEM touch/attach CT.LIQUID-ink
'This shirt has got ink on it.'
(b) *namø-mùk2 tit2 sùa4 nii4*
CT.LIQUID-ink touch/attach shirt DEM
'Ink has got on this shirt.'

(18) *man2 bang3 hùan2*
3SG.B block.from.view house
i. 'He's blocked from view by the house.'
ii. 'He's blocking the house from view.'

Alternations such as (16–18) are conditioned primarily by information structural considerations (i.e. responsive to dimensions of focus, topic, presupposition, as roughly captured in the different English translations). These are good illustrations

of the lack of relation between any available form of explicit coding (i.e. constituent order) and any particular type of grammatical relation (e.g. subject, object) or semantic role (e.g. theme, location).

57.3.3 Asymmetric bivalent predicates

There is greater variety in asymmetric bivalent relations, which map two arguments onto a predicate where the argument asymmetry has a straightforward directionality:

(19) *Transitive relation* (A=agent/effector, O=patient/theme)
Expresses the meaning 'A does V to O (which causes O to be in a certain state)' (e.g. *tom4* 'boil', *pia3* 'plat', *khaa5* 'kill', *puk2* 'waken').

(20) *Experiencer subject relation* (A=experiencer, O=theme)
Expresses the meaning 'A has the experience of V as a result of the stimulus O'; includes 'applied stimulus' expressions (e.g. *sèèp4* '(find something) delicious', *nak2* '(find something) heavy', *tùùn1* 'be startled (by something)'). There is an animacy constraint on the A.

(21) *Caused state relation* (A=effector, O=theme)
Expresses the meaning 'O comes to be in state V because of A'; includes 'caused state' expressions (e.g. *laaj2* '((cause to) become) striped', *dam3* '((cause to) become) black', *hòòn4* '((cause to) become) hot'). (Usually not agentive, although there are exceptions; e.g. *qun1* 'to warm something up'.)

(22) *Applied effector relation* (A=theme, O=effector)
Expresses the meaning 'A is in state V because of O'; includes (e.g. *vaan3* 'be sweet (because of something, e.g. sugar)', *phêt2* 'be spicy (because of something, e.g. chili)', *taaj3* 'die (because of something, e.g. sunlight)').

These patterns are interrelated in various ways. (The details are beyond our present scope.) While a few Lao verbs are relatively restricted in their argument structure (e.g. *tèèk5* 'to be broken, to break [intr.]' and *fot2* 'to be boiling' are both strictly monovalent), most verbs are versatile. To take one example, the verb *nak2* 'heavy' appears in four of the expression types listed above, one exception being the transitive relation:

(23) *kapaw3 nii4 nak2*
bag DEM heavy
'This bag is heavy.' (Inchoative-stative monovalent)

(24) *khòòj5 nak2 kapaw3 nii4*
1SG.POL heavy bag DEM
'I find this bag heavy.' (Experiencer subject)

978–0–19–920647–6 58-Spencer-ch57 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 816 of 820 May 16, 2008 19:1

(25) *kapaw3 nii4 nak2 law5*
bag DEM heavy liquor
'The bag is heavy from the liquor (inside it).' (Applied effector)

(26) *law5 nii4 nak2 kapaw3*
liquor DEM heavy bag
'This liquor makes the bag heavy.' (Caused state)

With the ever-present possibility of ellipsis in Lao, multiple interpretations become more likely. Thus, *khòòj5 nak2* [I heavy] could be a monovalent expression meaning 'I'm heavy' or an experiencer subject expression meaning 'I'm finding (it) heavy' (i.e. where the object argument is ellipsed). Only context will tell which it is.

A further confounding parameter for referential interpretation is 'ambi-valency', that is the possibility of a predicate entering into either a monovalent or a bivalent relation. The correct referential analysis (i.e. the one intended by the speaker) is only resolved by consulting the discourse record of a given usage:

(27) Surface sequence: **NP $V_{ambivalent}$**
Structural analysis 1: **NP_S V**
Structural analysis 2: **NP_A V** (NP_O ellipsed)
Structural analysis 3: **NP_O V** (NP_A ellipsed)
e.g.
(a) *paa3 kin3 lèèw4*
fish eat PRF
i. 'The fish has been eaten.'
ii. 'The fish has eaten (it).'
iii. 'The fish, (they) have eaten.' (constructed example; cf. Chao 1968: 75)
(b) *khèèw5 bòø than2 mii2*
tooth NEG be.on.time have/there.is
i. 'There were not yet any teeth.' (possible reading)
ii. 'The teeth didn't yet have (it/them).' (possible reading)
iii. 'Teeth, (it/they) didn't yet have.' (actual reading)

57.4 Expressive case functions: construal in event representation by marked construction

We have so far mostly been concerned with the disambiguating function of core case-marking – that is, the function of linking arguments to roles and/or indices. We

now consider the expressive functions which case-marking might perform, that is where special treatment of one or another argument serves to manipulate semantic distinctions in the construal of event structure (e.g. more versus less complete, aspectually), and participant involvement (e.g. more versus less involved, more versus less responsible). These kinds of distinction relate to transitivity in the sense of Hopper and Thompson (1980).

A key example in Lao is the 'handling-verb construction' (Enfield in press b), illustrated in (28b) as an alternative rendition of (28a).

(28) (a) *man2 thim5 pùm4*
3SG.B discard book
'He discards the book.' (Transitive construction)
(b) *man2 qaw3 pùm4 thim5*
3SG.B take book discard
'He takes the book (and) discards (it).' (Handling-verb construction)

Both (28a) and (28b) involve the same two arguments ('he', 'book'), and the same bivalent transitive verb ('discard'). The difference is that in (28b), the undergoer is expressed as a direct complement of an added verb of manipulation, with at least two effects. The first effect of this extra verb is to construe the event as bifurcated, breaking it down into two phases: 'coming into control of the undergoer' followed by 'despatch of the undergoer'. (In (28a), by contrast, the actor's prior control over the undergoer is presupposed.) The second effect is to change the relative order of the undergoer and the main content verb: in the marked construction, the undergoer is brought forward (if expressed; cf. (29), below), and the verb goes to a clause-final position. Much has been made of the possible long-term historical effects of such a construction taking hold in Sinitic languages such as Mandarin (e.g. the creation of object case-marking, and associated change of word order from SVO to SOV; Li and Thompson 1981: 26, 463ff).

It has been said in the Sinitic context that the added element which hosts the undergoer (here the verb 'take') is equivalent to a case marker, explicitly marking the semantic role of the undergoer (e.g. theme). In Lao it cannot be regarded as a case marker, for at least two reasons. First, the element is not an affix or other bound or dependent morpheme type. The item *qaw3* 'take' is a regular verb with regular verb trappings. For instance, it may ellipse its arguments if they are contextually retrievable, thus appearing with no dependent material whatsoever, as in this recasting of (28b):

(29) *qaw3 thim5*
take discard
'(He) takes (it and) discards (it).'

Second, 'object marking' by means of the handling-verb construction is not obligatory. It is a marked alternative to a single-verb transitive construction, whether

978–0–19–920647–6 58-Spencer-ch57 OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 818 of 820 May 16, 2008 19:1

this be an AVO structure as in (28a), above, or some version of it (with movement and/or ellipsis), as in the following:

(30) (a) *pùm4 man2 thim5*
book 3SG.B discard
'The book, he discards.'
(b) *thim5*
discard
'(He) discards (it).' (or: '(It) is discarded.')

If the function of the Lao handling-verb construction is at all comparable to that of case-marking, it is an expressive function, not a disambiguating function.

The handling-verb construction represents one possibility in an isolating language such as Lao to achieve an equivalent of one type of case-marking function, that is by singling out an argument for some sort of special treatment, as a way of manipulating the understood construal of the event predicated, where the relevant parameters of meaning typically relate to some or other of the ensemble of transitivity features (Hopper and Thompson 1980), including 'definiteness' and 'control'.

57.5 Conclusion

The formal organization of predicate–argument relations in a radically isolating language illustrates that the disambiguating functions of core case-marking need not be marked at all. In lieu of such case-marking, or any other unequivocal form of argument–role marking (agreement marking, strict/unambiguous interpretation of word order patterns), reference is well handled by the sheer pragmatics of context. The key source of information may be in the linguistic context (a hearer's constrained expectations about role and reference given selectional restrictions of predicates), or may be in the discourse/situational context (given the facts of particular discourse trajectories on particular occasions of language use). Grammarians of such languages have noted these problems, for example Thompson writing on Vietnamese: 'the familiar dichotomy of English verbs between those which "take objects" and those which do not is absent' (Thompson 1987: 220). Or as Gil puts it, writing on Riau Indonesian, languages of this kind are 'without distinct construction-specific semantic rules, compositional semantics relying instead on the association operator, which says that the meaning of a composite expression is associated with the meanings of its constituents in an underspecified fashion' (Gil 2005a: 1). Core grammatical relations are open to construal according to context. In actual language usage, rampant argument–role–reference ambiguity seldom poses communicative problems.

Like speakers of other languages, Lao speakers have grammatical resources for the optional expression of special construal of event–participant relations, similar to the kinds of function which might be performed by certain uses of case-marking in other languages. To reverse the perspective, such expressive functions might not be central functions of case anyway. That is to say, an expressive case-marking alternation (e.g. Finnish partitive marking on an undergoer) might just as well be viewed as a functional equivalent to a constructional alternation in a language like Lao. Perhaps when case-marking performs such expressive functions, it, too, is being appropriated, extended beyond its merely disambiguating prototype function to meet more expressive communicative goals.

References

Abaev, V. I. 1964. *A Grammatical Sketch of Ossetic*. Bloomington/The Hague: Mouton.

Abbott, B. 2004. Definiteness and Indefiniteness. In R. Horn and G. Ward (eds.). *Handbook of Pragmatics*. Oxford: Blackwell. 122–49.

Abboud, P. F. and E. N. McCarus (eds.). 1983. *Elementary Modern Standard Arabic, Parts 1, 2*. Cambridge: Cambridge University Press.

Abdullaev, S. 1954. *Grammatika darginskogo jazyka (fonetika i morfologija)*. Makhachkala: IIJaL.

Abney, S. 1987. *The English Noun Phrase in Its Sentential Aspect*. PhD diss., MIT, Cambridge, MA.

Abondolo, D. 1998. Hungarian. In D. Abondolo (ed.). *The Uralic Languages*. London: Routledge, 428–56.

Abraham, W. 2003. The Myth of Doubly Governing Prepositions in German. In E. Shay and U. Seibert (eds.). *Motion, Direction and Location in Languages*. Amsterdam/Philadelphia: John Benjamins. 19–38.

Ackerman, F. 1998. Constructions and Mixed Categories: Determining the Semantic Interpretation of Person/Number Marking. In M. Butt and T. H. King (eds.). *Proceedings of the LFG98 Conference*. Stanford, CA: CSLI Publications. [http://cslipublications.stanford.edu/LFG/3/lfg98.html]

—— and J. Moore 2001. *Proto-properties and Grammatical Encoding: A Correspondence Theory of Argument Selection*. Stanford: C.S.L.I.

Adelaar, W. F. H. and P. C. Muysken 2004. *The Languages of the Andes*. Cambridge: Cambridge University Press.

Aikhenvald, A. Y. 1990. On Berber Case in the Light of Afroasiatic Languages. In H. G. Mukarovsky (ed.). *Proceedings of the Fifth Hamito-Semitic Congress*, 1. Beiträge zur Afrikanistik, 40. Vienna: Institut für Afrikanistik und Ägyptologie der Universität Wien. 113–19.

—— 2000. *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

—— 2001. Mechanisms of Change in Areal Diffusion: New Morphology and Language Contact. In *Journal of Linguistics* 39. 1–29.

—— 2002. *Language Contact in Amazonia*. New York: Oxford University Press.

—— 2003. *A Grammar of Tariana from Northwest Amazonia*. Cambridge: Cambridge University Press.

—— R. M. W. Dixon, and M. Onishi (eds.). 2001. *Non-Canonical Marking of Subjects and Objects*. Amsterdam: John Benjamins.

Aissen, J. 1999. Markedness and Subject Choice in Optimality Theory. In *Natural Language and Linguistic Theory* 17. 673–711.

Aissen, J. 2003. Differential Object Marking: Iconicity vs. Economy. In *Natural Language and Linguistic Theory* 21. 435–83.

Akhtar, N., and M. Tomasello. 1997. 'Young Children's Productivity with Word Order and Verb Morphology'. *Developmental Psychology* 33: 952–65.

Akhutina, T. V. 1991. Is Agrammatism an Anomaly? In *Grazer Linguistische Studie* 35. 65–81.

Aksu-Koç, A. and N. Ketrez. 2007. Early Nominal Morphology in Turkish: Emergence of Case and Number. In U. Stephany and M. Voeikova (eds.). *Development of Nominal Inflection in First Language Acquisition: a Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter.

Aldridge, E. 2004. *Ergativity and Word Order in Austronesian Languages*. PhD thesis, Cornell University.

Alekseev, M. E. 2003. *Sravnitel'no-istoričeskaja morfologija naxsko-dagestanskix jazykov. Kategorii imeni*. Moscow: Academia.

—— and Ataev, B. M. 1998. *Avarskij jazyk*. Moscow: Academia.

Allan, K. 2007. *The Western Classical Tradition in Linguistics*. London: Equinox.

Allen, C. L. 1995. *Case Marking and Reanalysis: Grammatical Relations from Old to Early Modern English*. Oxford: Oxford University Press.

Allen, J. 1987. *Natural Language Understanding*. Menlo Park, CA: Benjamin/Cummings.

Alsina, A. 1996. The Role of Argument Structure in Grammar: Evidence from Romance. Stanford. CA: CSLI Publications.

Amberber, M. 1996. *Transitivity Alternations, Event-types and Light Verbs*. Unpublished doctoral dissertation, McGill University, Montreal.

—— 2000. Valency-Changing and Valency-Encoding Devices in Amharic. In R. M. W. Dixon and A. Y. Aikhenvald (eds.). *Changing Valency: Case Studies in Transitivity*. Cambridge: Cambridge University Press. 312–32.

—— 2001. Testing Emotional Universals in Amharic. In J. Harkins and A. Wierzbicka (eds.). *Emotions in Crosslinguistic Perspective*. Berlin and New York: Mouton de Gruyter. 39–72.

—— 2002. *Verb Classes and Tansitivity in Amharic*. Lincom Sudies in Afro-Asiatic Linguistics, 6. Munich: Lincom.

—— 2005. Differential Subject Marking in Amharic. In M. Amberber and H. de Hoop (eds.). *Competition and Variation in Natural Languages: the Case for Case*. Oxford: Elsevier. 295–321.

—— In press. The lexical Exponents of Semantic Primes in Amharic. In Cliff Goddard (ed.). *Crosslinguistic Semantics*. Amsterdam: John Benjamins.

Ambrazas, V. et al. 1985. *Grammatika litovskogo jazyka*. Vil'njus: Mokslas. [Engl. transl.: Ambrazas, V. et al. *Lithuanian Grammar*. Vilnius: Baltos lankos, 1997.]

Ameka, F. 1994. Ewe. In: Cliff Goddard and Anna Wierzbicka. (eds.). *Semantic and Lexical Universals: Theory and empirical findings.* Amsterdam/Philadelphia. John Benjamins, 57–86.

Amha, A. 2001. *A Grammar of Maale*. Dordrecht: Foris.

Andersen, T. 1988. Ergativity in Päri, a Nilotic OVS language. In *Lingua* 75. 289–324.

—— 2000. Anywa and Päri, II: A Morphosyntactic Comparison. In *Afrika und Übersee* 83. 65–87.

—— 2002. Case Inflection and Nominal Head Marking in Dinka. In *Journal of African Languages and Linguistics* 23. 1–30.

Anderson, J. M. 1971. *The Grammar of Case: Towards a Localistic Theory*. Cambridge: Cambridge University Press.

——1977. *On Case Grammar: Prolegomena to a Theory of Grammatical Relations*. London: Croom Helm.
——1984. Objecthood. In F. Plank (ed.). *Objects: Towards a Theory of Grammatical Relations*. New York & London: Academic Press. 29–54.
——1992. *Linguistic Representation: Structural Analogy and Stratification*. Berlin: Mouton de Gruyter.
——1997. *A Notional Theory of Syntactic Categories*. Cambridge: Cambridge University Press.
——2005. Structuralism and Autonomy: From Saussure to Chomsky. In *Historiographia Linguistica* 32.1/2. 117–48.
——2006. *Modern Grammars of Case: A Retrospective*. Oxford: Oxford University Press.
——2007. *The Grammar of Names*. Oxford: Oxford University Press.
Anderson, L. B. 1982. The 'perfect' as a Universal and as a Language-Particular Category. In P. Hopper (ed.). *Tense-Aspect: Between Semantics and Pragmatics*. Amsterdam: John Benjamins. 227–64.
Anderson, Lloyd B. 1986. Evidentials, Paths of Change, and Mental Maps: Typologically Regular Asymmetries. In W. Chafe and J. Nichols (eds.). *Evidentiality: The Linguistic Encoding of Epistemology*. Norwood: Ablex. 273–312.
Anderson, N. and M. Wade. 1988. Ergativity and Control in Folopa. In *Language and Linguistics in Melanesia* 19. 1–16.
Anderson, S. R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
Anderson, S. 1976. On the Notion of Subject in Ergative Languages. In C. N. Li (ed.). *Subject and Topic*. New York: Academic Press. 1–23.
Andrews, A. D. 1976. The VP-Complement Analysis in Modern Icelandic. In *Proceedings of the 6th Meeting of the North Eastern Linguistic Society*. [Reprinted in J. Maling and A. Zaenen (eds.). 1990. *Syntax and Semantics 24: Modern Icelandic Syntax.* Academic Press Inc. 165–85]
——1985a. The Major Functions of the Noun Phrase. In T. Shopen (ed.). *Language Typology and Syntactic Description*, vol 1: *Clause Structure*. Cambridge: Cambridge University Press. 62–154.
——1985b. *Studies in the Syntax of Relative and Comparative Clauses*. New York: Garland Press.
——1996. Semantic Case-Stacking and Inside-Out Onification. In *Australian Journal of Linguistics* 16. 1–54.
——2001. Non-Canonical A/S Marking in Icelandic. In A. Y. Aikhenvald, R. M. W. Dixon, and M. Onishi (eds.). *Non-Canonical Marking of Subjects and Objects*. Amsterdam: John Benjamins. 85–111.
——(2007). The Major Functions of the Noun Phrase. In T. Shopen (ed.). *Language Typology and Syntactic Description*, vol. 1 (2nd edition). Cambridge: Cambridge University Press. 132–223.
Anttila, A. and V. Fong. 2000. The Partitive Constraint in Optimality Theory. In *Journal of Semantics* 17. 281–314.
Aoun, J., and D. Sportiche. 1983. On the Formal Theory of Government. In *The Linguistic Review* 2. 211–36.
Apel, J. and A. Melinger. 2005. *Structural Persistence in the Absence of Global Structural Parallelism*. Poster presented at the 11th Annual Conference on Architectures and Mechanisms for Language Processing, Ghent, Belgium.

Apel, J. and A. Melinger. In preparation. Structural Persistence from Syntactically Non-Identical Primes.

Arensen, J. 1982. *Murle Grammar*. Juba: Summer Institute of Linguistics.

Ariel, M. 2000. The Development of Person Agreement Markers: From Pronouns to Higher Accessibility Markers. In M. Barlow and S. Kemmer (eds.). *Usage-based Models of Language*. Stanford: CSLI Publications. 197–260.

Aristar, A.R. 1996. The Relationship between Dative and Locative: Kurylowicz's Argument from a Typological Perspective. In *Diachronica* 13. 207–324.

——1997. Marking and Hierarchy Types and the Grammaticalization of Case-Markers. In *Studies in Language* 21. 313–68.

Arka, I.W. 2003. *Balinese Morphosyntax: A Lexical-functional Approach*. Canberra: Pacific Linguistics.

Aronoff, M. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.

Asher, R. E. and T. C. Kumari. 1997. *Malayalam*. London: Routledge.

Ashton, E. 1944. *Swahili Grammar*. London: Longmans.

Asudeh, A. 2001. Linking, Optionality, and Ambiguity in Marathi: An Optimality Theory Analysis. In P. Sells (eds.). Formal and Empirical Issues in Optimality Theoretic Syntax. Stanford, CA: CSLI Publications. 257–312.

Austin, P. 1981. *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.

——1982. Transitivity and Cognate Objects in Australian languages. In P. J. Hopper and S. A. Thompson (eds.). *Studies in Transitivity, Syntax and Semantics*, Volume 15. New York: Academic Press. 37–47.

——1995. Double Case Marking in Kanyara and Mantharta Languages, Western Australia. In F. Plank (ed.). *Double Case: Agreement by Suffixaufnahme*. Oxford: OUP. 363–79.

Authier, G. (n.d.) *Elements de grammaire kryz*. Ms.

——To appear. *Eléments de la langue kryz, dialecte d'Alik*. Paris: Société de Linguistique de Paris.

Avrutin, S. 2006. Weak Syntax. In Y. Grodzinsky and K. Amunts (eds.). *Broca's Region*. Oxford: Oxford University Press.

Axvlediani, G. S. (ed.). 1963. *Grammatika osetinskogo jazyka*. Vol. 1. Ordžonikidze.

Aznárez Mauleón, M. 2005. Approximación desde el Metalenguaje Semántico Natural a la semántica y la pragmática de algunas expresiones de sinceridad del español actual. In M. C. Velarde, R. G. Ruiz, and Ó. L. Lamas (eds.). *Estudios sobre lo metalingüistico (en español)*. Frankfurt am Main: Peter Lang. 233–56.

Babby, L. 1980. *Existential Sentences and Negation in Russian*. Ann Arbor: Karoma.

Babyonyshev, M. 1993. The Acquisition of Russian Case. In C. Phillips (ed.). *Papers on Case and Agreement II*. MIT Working Papers in Linguistics 19. Cambridge, MA: MIT Press. 1–44.

Bacz, B. 1991. On the Meaning of the Dative in Polish. LACUS forum. 246–55.

Bader, M. and J. Bayer. 2006. *Case and Linking in Language Comprehension–Evidence from German*. Heidelberg: Springer.

——and M. Meng. 1999. Subject–Object Ambiguities in German Embedded Clauses: An Across-the-Board Comparison. In *Journal of Psycholinguistic Research* 28. 121–43.

Baerman, M. and D. Brown. 2005. Case Syncretism. In M. Haspelmath, M. Dryer, D. Gil and B. Comrie (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 118–21.

——— ——— and G. Corbett. 2005. *The Syntax-Morphology Interface: A Study of Syncretism*. Cambridge Studies in Linguistics 109. Cambridge: Cambridge University Press.

Baker, C. L. 1979. Syntactic Theory and the Projection Problem. In *Linguistic Inquiry* 10. 533–81.

Baker, M. C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: The University of Chicago Press.

——— 1997. Thematic Roles and Syntactic Structure. In L. Haegeman (ed.). *Elements of Grammar. Handbook of Generative Syntax*. Dordrecht: Kluwer. 73–137.

Balke, T. 1999. Kasus im Sumerischen. In *Sprachtypologie und Universalienforschung* 52.2. 118–29.

——— 2006. *Das sumerische Dimensionalkasussystem*. Münster: Ugarit-Verlag.

Bally, C. 1926. L'Expression des idées de sphère personelle et de solidarité dans les langues indo-européennes. In F. Fankhauser and J. Jud (eds.). *Festschrift Louis Gauchat*. Aarau: Sauerländer. 68–78.

Baltin, M. and L. Barrett. 2002. *The Null Content of Null Case.* Unpublished manuscript, New York University.

Barðdal, J. 2001. *Case in Icelandic–A Synchronic, Diachronic and Comparative Approach*. Lundastudier i Nordisk Språkvetenskap 57. Lund: Department of Scandinavian Languages.

——— 2004. The Semantics of the Impersonal Construction in Icelandic, German and Faroese: Beyond Thematic Roles. In W. Abraham (ed.). *Focus on Germanic Typology*. Berlin: Akademie Verlag. 105–37.

——— (forthcoming, a) Predicting the Productivity of Argument Structure Constructions. In *Berkeley Linguistic Society* 32(1).

——— (forthcoming, b) 'The Development of Case in Germanic', in J. Barðdal and S. Chelliah (eds.). *The Role of Semantics and Pragmatics in the Development of Case*. Amsterdam: Benjamins.

——— (forthcoming, c). *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: Benjamins.

Barker, C. 1995. *Possessive Descriptions*. Stanford, CA: CSLI Publications.

Barlow, M. and C. A. Ferguson (eds.). 1988. *Agreement in Natural Language*. Stanford, CA: Centre for the Study of Language and Information.

Baron, I., M. Herslund, and F. Sørensen (eds.). 2001. *Dimensions of Possession*. Amsterdam: John Benjamins.

Bartens, H.-H. 1989. *Lehrbuch der saamischen (lappischen) Sprache*. Hamburg: Helmut Buske Verlag.

Bary, C. and P. de Swart. 2005. Arguments against Arguments: Additional Accusatives in Latin and Ancient Greek. In J. Gervain (ed.). *Proceedings of the Tenth ESSLLI Student Session*. 12–24.

Bashir, E. 1999. The Urdu and Hindi Ergative Postposition *ne*: Its Changing Role in the Grammar. In R. Singh (ed.). *The Yearbook of South Asian Languages and Linguistics.* New Delhi: Sage Publications. 11-36.

Bassarak, A. 1988. Morphologisierungstendenzen im Türkischen. In C. Römer (ed.). *2. Jenaer Semantik-Syntax-Symposium*. Jena: Friedrich-Schiller-Universität.

Bates, E. and B. MacWhinney. 1989. Functionalism and the Competition Model. In B. MacWhinney, and E. Bates (eds.). *The Crosslinguistic Study of Sentence Processing*. Cambridge: Cambridge University Press. 3–76.

Bates, E., B. Wulfeck, and B. MacWhinney. 1991. Cross-linguistic Research in Aphasia: An Overview. In *Brain and Language* 41. 123–48.

Baugh, A. C. and T. Cable. 1993. *A History of the English Language*. London: Routledge.

Bayer, J. 2004. Non-nominative Subjects in Comparison. In P. Bhaskararao, and K. V. Subbaraou (eds.). *Non-Nominative Subjects*, vol. 1. Amsterdam: John Benjamins. 49–76.

—— and J. Kornfilt. 1994. Against Scrambling as an Instance of Move-Alpha. In N. Corver and H. van Riemsdijk (eds.). *Studies on Scrambling; Movement and Non-Movement Approaches to Free Word-Order Phenomena*. Berline: Mouton de Gruyter. 17–60.

—— R. De Bleser, and C. Dronsek. 1987. Form und Funktion von Kasus bei Agrammatismus. In J. Bayer (ed.). *Grammatik und Kognition*. Linguistische Berichte Sonderheft 1. Wiesbaden: Westdeutscher Verlag. 81–117.

—— M. Bader, and M. Meng. (2001). Morphological Underspecification Meets Oblique Case: Syntactic and Processing Effects in German. *Lingua* 111. 465–514.

Beard, R. 1995. *Lexeme Morpheme Base Morphology*. Stony Brook, NY: SUNY Press.

Beekes, E. S. P. 1995. *Comparative Indo-European Linguistics: An Introduction*. Amsterdam: John Benjamins.

van Belle, W. and W. van Langendonck. 1996. *The Dative*. Volume 1: *Descriptive Studies*. Amsterdam/Philadelphia: John Benjamins.

Belyavski-Frank, M. 1983. The Pattern of some (manly) Morphological Features in Southern Serbian Dialects. In *Makedonski jazik*. Kn. 34. Skopje. 213–30.

Bennett, P. R. 1974. Tone and the Nilotic Case System. In *Bulletin of the School of Oriental and African Studies* 37(1). 19–28.

Benzing, J. 1955. *Lamutische Grammatik*. Wiesbaden: Franz Steiner.

—— 1985. *Kalmückische Grammatik zum Nachschlagen*. Wiesbaden: Harrassowitz.

Berg, T. 1987. The Case Against Accommodation: Evidence from German Speech Error Data. In *Journal of Memory and Language* 26. 277–99.

van den Berg, H. E. 1995. *A Grammar of Hunzib* (with Texts and Lexicon). Munich/Newcastle: Lincom Europa.

—— 2001. *Dargi Folktales. Oral Stories From the Caucasus with an Introduction to Dargi Grammar*. Leiden: CNWS.

van den Berg, H. 2005. The East Caucasian Language Family. In *Lingua* 115. 147–90.

Bergsland, K. 1997. *Aleut Grammar: Unangam Tunuganaan Achixaasix.* Fairbanks: ANLC (Alaska Native Language Center).

Berndt, R. S., C. C. Mitchum, and A. N. Haendiges. 1996. Comprehension of Reversible Sentences in 'Agrammatism': A Meta-analysis. In *Cognition* 58. 289–308.

Bernhardi, A.-F. 1805. *Anfangsgründe der Sprachwissenschaft*. Berlin.

Berry, K. and C. Berry. 1999. *A Description of Abun: A West Papuan Language of Irian Jaya*. Pacific Linguistics B, 115. Canberra, Australia: Research School of Pacific and Asian Studies, Australian National University.

Bertolo, S. (ed.) 2001. *Language Acquisition and Learnability*. Cambridge: Cambridge University Press.

Bhaskararao, P. and K.V. Subbarao (eds.). 2004. *Non-nominative Subjects* (Vols. 1 & 2). Amsterdam: John Benjamins.

Bhat, D. N. S. 1994. *The Adjectival Category*. Amsterdam/Philadelphia: John Benjamins.

—— 2004. *Pronouns*. Oxford/New York: Oxford University Press.

—— and M. S. Ningomba. 1997. *Manipuri Grammar*. Munich/Newcastle: Lincon.

Bickel, B. 2003. Belhare. In G. Thurgood and R. J. LaPolla (eds.). *The Sino-Tibetan Languages*. London: Routledge. 546–70.

——2004a. Hidden Syntax in Belhare. A. Saxena (ed.). *Himalayan Languages: Past and Present*. Berlin: Mouton de Gruyter. 141–90.

——2004b. The Syntax of Experiencers in the Himalayas. In P. Bhaskararao and K. V. Subbarao (eds.). *Non-nominative subjects*, Vol. 1. Amsterdam: Benjamins. 77–113.

——2006. Clause-Level vs. Predicate-Level Linking. In I. Bornkessel, M. Schlesewsky, B. Comrie, and A. D. Friederici (eds.). *Semantic Role Universals and Argument Linking.* 155–90. Berlin: Mouton de Gruyter.

——2007a. Grammatical Relations Typology. In J. J. Song (ed.). *The Oxford Handbook of Language Typology*. Oxford: Oxford University Press.

——2007b. Typology in the 21st Century: Major Current Developments. In *Linguistic Typology* 11. 239–51.

——(forthcoming, a). On the Scope of the Referential Hierarchy in the Typology of Grammatical Relations. In G. G. Corbett and M. Noonan (eds.). *Case and Grammatical Relations: Papers in Hanour of Bernand Comrie*. Amsterdam: John Benjamins.

——(forthcoming, b). A Refined Sampling Procedure for Genealogical Control. In *Sprachtypologie und Universalienforschung*.

——and J. Nichols. 1996ff. *The AUTOTYP database*. Electronic database: http://www.uni-leipzig.de/~autotyp.

————2003. *Typological Enclaves*. Paper presented at the 5th Biannual Conference of the Association for Linguistic Typology, Cagliari, September 18. Available at http://www.uni-leipzig.de/~autotyp/download.

————2005a. Exponence of Selected Inflectional Formatives. In M. Haspelmath et al. (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 90–93.

————2005b. Inclusive/Exclusive as Person vs. Number Categories Worldwide. In E. Filimonova (ed.). *Clusivity*. Amsterdam: Benjamins. 47–70.

————2005c. Fusion of Selected Inflectional Formatives. In M. Haspelmath et al. (eds.). *World Atlas of Language Structures*. Oxford: Oxford University Press. 86–9.

————2006. Oceania, the Pacific Rim, and the Theory of Linguistic Areas. In *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society*. [PDF available at http://www.uni-leipzig.de/~bickel/research/papers.]

————2007. Inflectional Morphology. In T. Shopen (ed.). *Language Typology and Syntactic Description* (Revised second edition). Cambridge: Cambridge University Press.

————(forthcoming). *Generalizing Alignment Typology*. Ms. in preparation.

Bierwisch, M. 1988. On the Grammar of Local Prepositions. In M. Bierwisch et al. (eds.). *Syntax, Semantik und Lexikon*. Berlin: Akademie. 1–66.

Birnbaum, H. 1965. Balkanslavisch und Südslavisch. In *Zeitschrift für Balkanologie*. Jg. III. Wiesbaden. 12–63.

Bisang, W. 1992. *Das Verb im Chinesischen, Hmong, Vietnamesischen, Thai und Khmer. Vergleichende Grammatik im Rahmen der Verbserialisierung, der Grammatikalisierung und der Attraktorpositionen*. Tübingen: Narr.

Bittner, D., W. U. Dressler, and M. Kilani-Schoch (eds.). 2000. *First Verbs: On the Way to Mini-paradigms*. ZAS Papers in Linguistics 18. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung.

————(eds.). 2003. *Development of Verb Inflection in First Language Acquisition: A Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter.

Bittner, M. and K. Hale. 1996a. The Structural Determination of Case and Agreement. In *Linguistic Inquiry* 27. 1–68.

——— 1996b. Ergativity: Toward a Theory of a Heterogeneous Class. In *Linguistic Inquiry* 27. 531–604.

Blake, B. J. 1976. On Ergativity and the Notion of Subject. In *Lingua* 39. 281–300.

—— 1977. *Case Marking in Australian Languages*. AIAS Linguistic Series 23. Canberra: Australian Institute of Aboriginal Studies.

—— 1994. *Case*. Cambridge Textbooks in Linguistics 32. Cambridge: Cambridge University Press.

—— 2001. *Case* (2nd edition). Cambridge: Cambridge University Press.

Blanchon, J. A. 1998. Semantic/Pragmatic Conditions on the Tonology of the Kongo Noun-Phrase: A Diachronic Hypothesis. In L. Hyman and C. Kisseberth (eds.). *Theoretical Aspects of Bantu Tone*. Stanford: Stanford University Press. (Center for the Study of Language and Information.) 1–32.

—— 1999. Tone Cases in Bantu Group B.40. In J. A. Blanchon, and D. Creissels (eds.). *Issues in Bantu Tonology*. Cologne: Rüdiger Köppe. 37–82.

Blank, A. 2004. Polysemy in the Lexicon and in Discourse. 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. 267–93.

Blansitt, E. L. Jr. 1973. *Bitransitive Clauses*. Working Papers on Language Universals.

—— 1988. Datives and Allatives. In M. Hammond, E. Moravcsik, and J. Wirth (eds.). *Studies in Syntactic Typology*. Amsterdam/Philadelphia: John Benjamins. 173–91.

De Bleser, R., C. Dronsek, and J. Bayer. 1988. Morphosyntactic Processing in German Agrammatism: A Replication and Revision of von Stockert/Bader (1976). In *Cortex* 24. 53–76.

—— F. Burchert, and P. Rausch. 2005. Breakdown at the Morphological Level in Agrammatism. In *Stem- Spraak- Taalpathologie* 13. 35–46.

Blevins, J. P. 2005. Word-based Declensions in Estonian. In G. Booij and J. van Marle (eds.). *Yearbook of Morphology 2005*. Dordrecht: Springer. 1–25.

—— 2006. Word-based Morphology. In *Journal of Linguistics* 42. 531–73.

Bloch, J. 1934. *L'indo-aryen: du véda aux temps modernes*. Paris: Adrien-Maisonneuve.

Blumstein, S. E., G. Byma, K. Kurowski, J. Hourihan, T. Brown, and A. Hutchinson. 1998. On-line Processing of Filler-Gap Constructions in Aphasia. In *Brain and Language* 61(2). 149–68.

Blutner, R. 2000. Some Aspects of Optimality in Natural Language Interpretation. *Journal of Semantics* 17. 189–216.

Boas, F. 1964 [1911]. On Grammatical Categories. In D. Hymes (ed.). *Language in Culture and Society: A Reader in Linguistics and Anthropology*. New York: Harper & Row. 121–3.

Bobaljik, J. D. 1993a. On Ergativity and Ergative Unergatives. In C. Phillips (ed.). *Papers on Case and Agreement II*. MIT Working Papers in Linguistics 19. 45–88.

—— 1993b. Nominally Absolutive is not Absolutely Nominative. In *Proceedings of the Eleventh West Coast Conference on Formal Linguistics*. University of California, Stanford: CLSI.

—— (to appear). Where's phi? Agreement as a Post-syntactic Operation. In D. Harbour, D. Adger, and S. Béjar (eds.). *Phi-Theory: Phi Features across Interfaces and Modules*. Oxford: Oxford University Press.

——and P. Branigan. 2006. Eccentric Agreement and Multiple Case Checking. In A. Johns, D. Massam, and J. Ndayiragije (eds.). *Ergativity: Emerging Issues*. Dordrecht: Springer. 47–77.

Bock, J. K. 1986. Meaning, Sound, and Syntax: Lexical Priming in Sentence Production. In *Journal of Experimental Psychology: Learning, Memory and Cognition* 12. 575–86.

——1987. Co-ordinating Words and Syntax in Speech Plans. In A. W. Ellis (ed.). *Progress in the Psychology of Language*, Vol. 3. Hillsdale, NJ: Erlbaum. 337–90.

——1989. Closed-class Immanence in Sentence Production. In *Cognition* 31. 163–86.

——1995. Sentence Production: From Mind to Mouth. In J. L. Miller and P. D. Eimas (eds.). *Handbook of Perception and Cognition: Speech, Language, and Communication*, Vol. 11. Orlando, FL: Academic Press. 181–216.

——and Z. M. Griffin. 2000. The Persistence of Structural Priming: Transient Activation or Implicit Learning? In *Journal of Experimental Psychology: General* 129. 177–92.

——and W. J. M. Levelt. 1994. Language Production: Grammatical Encoding. In M. A. Gernsbacher (Ed.). *Handbook of Psycholinguistics*. San Diego, CA: Academic Press. 945–84.

——and H. Loebell. 1990. Framing Sentences. In *Cognition* 35. 1–39.

—— ——and R. Morey. 1992. From Conceptual Roles to Structural Relations: Bridging the Syntactic Cleft. *Psychological Review* 99. 150–71.

Boeder, W. 1979. Ergative Syntax and Morphology in Language Change: The South Caucasian Languages. In F. Plank (ed.). *Ergativity: Towards a Theory of Grammatical Relations*. London: Academic Press. 435–80.

——1995. Suffixaufnahme in Kartvelian. In F. Plank (ed.). *Double Case. Agreement by Suffixaufnahme*. Oxford: Oxford University Press, 151–215.

Bohas, G, J.-P. Guillaume, and D.E. Kouloughli. 1990. *The Arabic Linguistic Tradition*. London/New York: Routledge

Böhm, R. 1982. *Topics in Localist Case Grammar (with special reference to English and German)*. DPhil thesis, New University of Ulster.

——1998. *Notional Grammar, Wortklassen und Dependenz*. (BLicK 7) Bremen: Universität Bremen, IAAS.

——1999. Kategorien, Konfigurationen und 'Kasus' in einer notionalen Grammatik. In *Sprachtypologie und Universalienforschung* 52.2. 207–31.

Bokarev, A. A. 1949a. *Sintaksis avarskogo jazyka*. Moscow/Leningrad: AN SSSR.

——1949b. *Ocherki grammatiki chamalinskogo jazyka*. Moscow: AN SSSR.

Borer, H. 1989. On the Morphological Parallelism between Compounds and Constructs. In *Yearbook of Morphology I*. Dordrecht: Kluwer. 45–65.

——2005. *Structuring Sense*. Oxford: Oxford University Press.

Borg, A. and M. Azzopardi-Alexander. 1995. *Maltese*. London: Routledge.

Borgman, D. M. 1990. Sanuma. In D. C. Derbyshire and G. K. Pullum (eds.). *Handbook of Amazonian Languages*, vol. 2. Berlin: Mouton de Gruyter. 15–248.

Bornkessel, I. and M. Schlesewsky. 2006. The Extended Argument Dependency Model: A Neurocognitive Approach to Sentence Comprehension Across Languages. In *Psychological Review* 113. 787–821.

—— ——and A. D. Friederici. 2002. Grammar Overrides Frequency: Evidence from the Online Processing of Flexible Word Order. In *Cognition* 85(2). B21–B30.

Bornkessel, I., M. Schlesewsky, and A. D. Friederici. 2003. Eliciting Thematic Reanalysis Effects: The Role of Syntax-Independent Information During Parsing. In *Language and Cognitive Processes* 18. 269–98.

—— B. McElree, M. Schlesewsky, and A. D. Friederici. (2004). Multi-dimensional Contributions to Garden Path Strength: Dissociating Phrase Structure from Case Marking. In *Journal of Memory and Language* 51: 495–522.

Bošković, Ž. 1996. Selection and the Categorial Status of Infinitival Complements. In *Natural Language and Linguistic Theory* 14. 269–304.

—— 1997. *The Syntax of Nonfinite Complementation: An Economy Approach.* Cambridge, MA: MIT Press.

—— 2002. A-Movement and the EPP. In *Syntax* 5. 167–218.

—— 2005. On the Locality of Left Branch Extraction and the Structure of NP. In *Studia Linguistica* 59(1). 1–45.

—— (to appear). On the Locality and Motivation of Move and Agree: an Even More Minimal Theory. In *Linguistic Inquiry.*

Bossong, G. 1985a. *Differenzielle Objektmarkierung in den neuiranischen Sprachen.* Tübingen: Narr.

—— 1985b. Markierung von Aktantenfunktionen in Guarani. Zur Frage der differenziellen Objektmarkierung in nicht-akkusativischen Sprachen. In F. Plank (ed.). *Relational Typology*. Berlin: Mouton. 1–31.

—— 1985c. Zur Entwicklungsdynamik von Kasussystemen. In *Folia Linguistica Historica* 6, 2. 285–321.

—— 1985d. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Ars linguistica. Commentationes analyticae et criticae 14. Tübingen: Gunther Narr Verlag.

—— 1991. Differential Object Marking in Romance and Beyond. In D. Wanner and D. A. Kibbee (eds.). *New Analyses in Romance Linguistics: Selected Papers from the XVIII Linguistics Symposium on Romance Languages, Urbana-Champaign, April 7–9, 1988*. Amsterdam: John Benjamins. 143–70.

—— 1998. Le marquage de l'expérient dans les langues d'Europe. In J. Feuillet (ed.). *Actance et valence dans les langues de l'Europe*. Berlin: Mouton de Gruyter. 259–94.

Bouchard, D. 1996. Sign Languages and Language Universals: The Status of Order and Position in Grammar. In *Sign Language Studies* 91. 101–60.

—— 1998. The Distribution and Interpretation of Adjectives in French: A Consequence of Bare Phrase Structure. In *Probus* 10. 139–83.

Bowe, H. 1990. *Categories, Constituents, and Constituent Order in Pitjantjatjara*. London/New York: Routledge.

Bowerman, M. 1985. What Shapes Children's Grammars? In D. I. Slobin (ed.). *The Crosslinguistic Study of Language Acquisition*. Volume 2. Hillsdale, NJ: Erlbaum. 1257–1314.

—— P. Brown, S. Eisenbeiss, B. Narasimhan, and D. Slobin. 2002. *Putting Things in Places: Developmental Consequences of Linguistic Typology*. Invited plenary symposium, Child Language Research Forum, Stanford University, April 2002. Online Proceedings: http://cslipublications.stanford.edu/CLRF/2002/CLRF-2002-title.html

Braine, M. D. S. 1976. *Children's First Word Combinations*. Monographs of the Society for Research in Child Development 41(1).

—— 1992. What Sort of Innate Structure is Needed to 'Bootstrap' into Syntax? In *Cognition* 45. 77–100.

Branigan, H. P., M. J. Pickering, and A. A. Cleland. 2000. Syntactic Coordination in dialogue. In *Cognition* 75. B13–B25.

————J. F. McLean, and A. Stewart. 2006. The Role of Local and Global Syntactic Structure in Language Production: Evidence from Syntactic Priming. *Language and Cognitive Processes* 21. 974–1010.

de Bray, R. G. A. 1980. *Guide to the West Slavonic Languages*. Columbus: Slavica.

Bresnan, J. 1977. Transformations and Categories in Syntax. In R. E. Butts and J. Hintikka (eds.). Basic Problems in Methodology and Linguistics. Part Three of the Proceedings of the Fifth International Congress of Logic, Methodology and Philosophy of Science, London, Ontario, Canada, 1975. Dordrecht: Reidel. 261–82.

——2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

——(ed.). 1982. *The Mental Representation of Grammatical Relations*. Cambridge: MIT Press.

——2000. Optimal Syntax. In J. Dekkers, F. van der Leeuw, and J. van de Weije (eds.). *Optimality Theory: Phonology, Syntax and Acquisition.* Oxford: Oxford University Press. 334–85.

——2006. Is Syntactic Knowledge Probabilistic? Experiments with the English Dative Alternation. Invited paper to appear in the proceedings of the International Conference on Linguistic Evidence, Tübingen, 2–4 February 2006, *Roots: Linguistics in Search of its Evidential Base*. Series: Studies in Generative Grammar, edited by Sam Featherston and Wolfgang Sternefeld. Berlin: Mouton de Gruyter.

——and J. Kanerva. 1989. Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. Linguistic Inquiry 20: 1–50.

——and R. M. Kaplan. 1982. Introduction: Grammars as Mental Representations of Language. In J. Bresnan (ed.). *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press. xvii–lii.

——and S. A. Mchombo. 1986. Grammatical and Anaphoric Agreement. In A. Farley, P. Farley, and K. McCullough (eds.). *Proceedings of the Chicago Linguistic Society 22. Papers from the Parasession on Pragmatics and Grammatical Theory*. 741–82.

Broadbent, S. 1964. *The Southern Sierra Miwok Language*. Berkeley: University of California Press.

Bromley, M. H. 1981. *A Grammar of Lower Grand Valley Dani*. Pacific Linguistic Series C. No.63. Canberra: Australian National University.

Brown, L. 2001. *A Grammar of Nias Selatan*. PhD Thesis, University of Sydney.

Brown, R. 1988. Waris Case System and Verb Classification. In *Language and Linguistics in Melanesia* 19. 37–80.

——and C. Hanlon. 1970. Derivational Complexity and Order of Acquisition in Child Speech. In J. R. Hayes (ed.). *Cognition and the Development of Language*. New York: Wiley. 11–53.

Bubeník, V. 1998. A *Historical Syntax of Late Middle Indo-Aryan (Apabhraüśa)*. Current Issues in Linguistic Theory 165. Amsterdam: Benjamins.

——(n.d.). *On the Evolution of the Hittite System of Postpositions from Proto-Indo-European*. Typescript.

——and J. Hewson. 2006. *From Case to Adposition: The Development of Configurational Syntax in Indo-European Languages*. Current Issues in Linguistic Theory 280. Amsterdam: Benjamins.

Buchholz, O. and W. Fiedler. 1987. *Albanische Grammatik*. Leipzig: VEB Verlag Enzyklopädie.

Budwig, N. 1989. 'The Linguistic Marking of Agentivity and Control in Child Language'. *Journal of Child Language* 16. 263–84.

Bugenhagen, R. 2002. The Syntax of Semantic Primes in Mangaaba-Mbula. In C. Goddard and A. Wierzbicka (eds.). *Meaning and Universal Grammar: Theory and Empirical Findings*. vol. 2. Amsterdam: John Benjamins. 1–64.

Bulut, C. 2006. Turkish Elements in Spoken Kurmanji. In H. Boeschoten and L. Johanson (eds.). *Turkic Languages in Contact*. Wiesbaden: Harrassowitz. 95–121.

Burchert, F., R. De Bleser, and K. Sonntag. 2001. Does Case Make the Difference? Agrammatic Comprehension of Case-inflected Sentences in German. In *Cortex* 37. 700–2.

——— ——— ——— 2003. Does Morphology Make the Difference? Agrammatic Sentence Comprehension in German. In *Brain and Language* 87. 323–42.

Burling, R. 2004. *The Language of the Modhupur Mandi (Garo)*. New Delhi: Bibliophile South Asia.

Burzio, L. 1981. *Intransitive Verbs and Auxiliaries*. Ph.D. Dissertation, MIT, Cambridge, MA.

Buth, R. 1981. Ergative Word Order–Luwo is OVS. In *Occasional Papers in the Study of Sudanese Languages* 1. 74–90.

Butt, M. 1995. *The Structure of Complex Predicates in Urdu*. Stanford: CSLI Publications.

——— 1998. Constraining Argument Merger through Aspect. In E. Hinrichs, A. Kathol, and T. Nakazawa (eds.). *Complex Predicates in Nonderivational Syntax*. New York: Academic Press. 73–113.

——— 2005. *Differential Case Marking Theory (DCT)*. (Ms). at http://www.hum.uit.no/a/svenonius/papers/Butt05CaseHO.pdf

——— 2006. *Theories of Case*. Cambridge: Cambridge University Press.

——— and T. H. King. 1991. Semantic Case in Urdu. In L. Dobrin, L. Nichols, and R. M. Rodriguez (eds.). *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*. 31–45.

——— ——— 2003. Case Systems: Beyond Structural Distinctions. In E. Brandner and H. Zinsmeister (eds.). *New Perspectives on Case Theory.* Stanford: CSLI Publications. 53–87.

——— ——— 2004. The Status of Case. In V. Dayal and A. Mahajan (eds.). *Clause Structure in South Asian languages*. Dordrecht: Kluwer. 153–98.

——— ——— and A. Varghese. 2004. A Computational Treatment of Differential Case Marking in Malayalam. In *Proceedings of the International Conference on Natural Language Processing (ICON) 2004*. Hyderabad.

Butterworth, B. 1980. Some constraints on Models of Language Production. In B. Butterworth (ed.). *Language Production*, Vol. 1: *Speech and Talk*. London: Academic Press. 423–59.

——— and D. Howard. 1987. Paragrammatisms. In *Cognition* 26. 2–37.

Bybee, J. L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.

——— 2006. *Frequency of Use and the Organisation of Language*. Oxford University Press.

——— R. D. Perkins, and W. Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Cadiot, P. 1990. La préposition avec: grammaire et représentation. In *Le Français Moderne* 58. 152–73.

Cain, B. D. and J. W. Gair. 2000. *Dhivehi (Maldivian)*. Munich: Lincom Europa.

Calabrese, A. 1998. Some Remarks on the Latin Case System and its Development in Romance. In J. Lema and E. Treviño (eds.). *Theoretical Analyses on Romance Languages*. Amsterdam: Benjamins. 71–126.

Caramazza, A. and R. S. Berndt. 1985. A Multicomponent Deficit View of Agrammatic Broca's Aphasia. In M. L. Kean (ed.). *Agrammatism*. Orlando: Academic Press. 27–63.

—— and E. Zurif. 1976. Dissociation of Algorithmic and Heuristic Processes in Language Comprehension: Evidence from Aphasia. In *Brain and Language* 3. 572–82.

—— C. Papagno, and W. Ruml. 2000. The Selective Impairment of Phonological Processing in Speech Production. In *Brain and Language* 75. 428–50.

Cardona, G. 1976a. Panini: A Survey of Research. In W. Winter (ed.). *Trends in Linguistics, State-of-the-Art Reports* (6). The Hague: Mouton.

——1976b. Subject in Sanskrit. In M. K. Verma (ed.). *The Notion of Subject in South Asian Languages*. Madison: University of Wisconsin, South Asian Studies, 1–38.

Carlson, G. 1977. A Unified Analysis of the English Bare Plural. In *Linguistics and Philosophy* 1. 413–58.

Carstairs, A. 1984. Outlines of a Constraint on Syncretism. In *Folia Linguistica* 18. 73–84.

——1987. *Allomorphy in Inflection*. London/New York/Sydney: Croom Helm.

Cecchetto, C. 2004. A Challenge to Null Case Theory. In *Linguistic Inquiry* 35.1. 141–49.

Ceytlin, S. N. 1989. *Child Speech: Innovations of Form and Word Building*. Unpublished doctoral dissertation, Leningrad.

Chadwick, N. 1976. Ergative, Locative and Instrumental Case Inflections–The Western Barkly Languages. In R. M. W. Dixon (ed.). *Grammatical Categories in Australian Languages*. Linguistic Series No. 22. Canberra: Australian Institute of Aboriginal Studies/New Jersey: Humanities Press. 390–6.

Chafe, W. 1971. *Meaning and the Structure of Language*. University of Chicago Press.

Chaker, S. 1983. *Un parler berbère d' Algérie (Kabylie). Syntaxe*. Aix-en-Provence: Publications de l'Université de Provence.

——1988. L'état d'annexion du nom. In *Encyclopédie berbère*, V. Paris: EDISUD, 686–95.

Chang, F. 2002. Symbolically Speaking: A Connectionist Model of Sentence Production. In *Cognitive Science* 26. 609–51.

—— G. S. Dell, K. Bock, and Z. M. Griffin. 2000. Structural Priming as Implicit Learning: A Comparison of Models of Sentence Production. In *Journal of Psycholinguistic Research* 29(2). 217–29.

—— —— ——2006. Becoming Syntactic. In *Psychological Review* 113, 2. 234–72.

Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. Berkeley/Los Angeles: University of California Press.

Chappell, H. 2002. The Universal Syntax of Semantic Primes in Mandarin Chinese. In C. Goddard and A. Wierzbicka (eds.). *Meaning and Universal Grammar: Theory and Empirical Findings*. vol. 1. Amsterdam: John Benjamins. 243–322.

—— and W. McGregor (eds.). 1995. *The Grammar of Inalienability: A Typological Perspective on Body Part Terms and the Part–Whole Relation*. Berlin: Mouton de Gruyter.

Charachidzé, G. 1981. *Grammaire de la langue avar (langue du Caucase Nord-Est)*. Paris: Editions Jean-Favard.

Charney, J. O. 1993. *A Grammar of Comanche*. Lincoln: University of Nebraska Press.

Chelliah, S. L. 1997. *A Grammar of Meithei*. Berlin: Mouton de Gruyter.

CHEUNG, J. (forthcoming). The Ossetic Case System Revisited. In A. Lubotsky, J. Schaeken, and J. Wiedenhof (eds.). *Evidence and Counter-evidence: Festschrift F. Kortlandt*. Vol. I (SSGL 33). Amsterdam: Rodopi.

CHOMSKY, N. 1957. *Syntactic Structures*. The Hague: Mouton de Gruyter.

——1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

——1966. *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York: Harper & Row.

——1980. On Binding. In *Linguistic Inquiry* 11. 1–46.

——1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.

——1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

——1986. *Barriers*. Cambridge, MA: MIT Press.

——1991. Some Notes on the Economy of Derivation and Representation. In R. Freidin (ed.). *Principles and Parameters in Comparative Grammar.* Cambridge, MA: MIT Press. 417–54.

——1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

——2000. Minimalist Inquiries: The Framework. In R. Martin, D. Michaels, and J. Uriagereka (eds.). *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press. 89–155.

——2001. Derivation by Phase. In M. Kenstowicz (ed.). *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press. 1–52.

——2004. *On Phases*. Ms., MIT.

—— and H. LASNIK. 1977. Filters and Control. In *Linguistic Inquiry* 8. 425–504.

—— ——1995. The Theory of Principles and Parameters. In N. Chomsky (ed.). *The Minimalist Program*. Cambridge, MA: MIT Press. 13–127

CHOUINARD, M. M. and E. V. CLARK. 2003. Adult Reformulations of Child Errors as Negative Evidence. In *Journal of Child Language* 30:3. 637–69.

CHRISTIANSON, K. and F. FERREIRA. 2005. Conceptual Accessibility and Sentence Production in a Free Word Order Language (Odawa). In *Cognition* 98. 105–35.

CHRISTENSEN, A. 1921. *Les dialectes d'Awromān et de Pāwä*. Copenhagen: Bianco Lunos Bogtrykkeri.

——1935. *Contributions à la dialectologie iranienne, II, Dialects de la région de Sèmnān: Sourkhéi, Lāsguerdi, Sängesäri et Chämerzådi*. Copenhagen: Levin & Munksgaard.

CHRISTOFIDOU, A., and U. STEPHANY. 1997. The Early Development of Case Forms in the Speech of a Greek Boy: A Preliminary Investigation. In K. Dziubalska-Kołaczyk (ed.). *Pre- and Protomorphology in Language Acquisition*. Papers and Studies in Contrastive Linguistics, V. 33. Poznań: Adam Mickiewicz University. 127–41.

CLAHSEN, H., S. EISENBEISS, and A. VAINIKKA. 1994. The Seeds of Structure. A Syntactic Analysis of the Acquisition of Case Marking. In T. Hoekstra and B. D. Schwartz (eds.). *Language Acquisition Studies in Generative Grammar*. Amsterdam: Benjamins. 85–118.

—— ——and M. PENKE. 1996. Lexical Learning in Early Syntactic Development. In H. Clahsen (ed.). *Generative Perspectives on Language Acquisition. Empirical Findings, Theoretical Considerations and Crosslinguistic Comparisons*. Amsterdam: Benjamins. 129–59.

CLELAND, S. and M. PICKERING. 2003. The Use of Lexical and Syntactic Information in Language Production: Evidence from the Priming of Noun-phrase Structure. In *Journal of Memory and Language* 49. 214–30.

Clifton, C., Jr. and L. Frazier. 1989. Comprehending Sentences with Long-Distance Dependencies. In G. N. Carlson and M. K. Tanenhaus (eds.). *Linguistic Structure in Language Processing*. Dordrecht: Kluwer. 273–317.

Clyne, M. 2003. *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: Cambridge University Press.

Colarusso, J. 1992. *A Grammar of the Kabardian Language*. Calgary: University of Calgary Press.

Cole, P. 1982. *Imbabura Quechua*. Lingua Descriptive Studies 5. Amsterdam: North Holland.

——1985. *Imbabura Quechua*. Amsterdam: North-Holland.

Collins, C. and H. Thráinsson. 1996. VP-Internal Structure and Object Shift in Icelandic. In *Linguistic Inquiry* 27. 391–444.

Comrie, B. 1976a. *Aspect*. Cambridge: Cambridge University Press.

——1976b. The Syntax of Causative Constructions: Cross-language Similarities and Divergences. In M. Shibatani (ed.). *Syntax and Semantics 6: The Grammar of Causative Constructions*. New York: Academic Press. 261–312.

——1978. Ergativity. In W. P. Lehmann (ed.). *Syntactic Typology. Studies in the Phenomenology of Language*. Austin, London: The University of Texas Press. 329–94.

——1980. The Order of Case and Possessive Suffixes in Uralic Languages: An Approach to the Comparative-historical Problem. In *Lingua Posnaniensis* 23. 81–6.

——1981a. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.

——1981b. (with B. G. Hewitt, J. R. Payne). *The Languages of the Soviet Union*. Cambridge: Cambridge University Press.

——1982. Grammatical Relations in Huichol. In P. J. Hopper and S. A. Thompson (eds.). *Syntax and Semantics 15: Studies in Transitivity*. New York/London: Academic Press. 95–115.

——1984. Case and Agreement in Hindi and Related Languages. In *Linguistics* 22. 857–64.

——1985. Causative Verb Formation and Other Verb-deriving Morphology. In T. Shopen (ed.). *Language Typology and Syntactic Description*, vol. III: *Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press. 309–48.

——1986. On Delimiting Cases. In R. D. Brecht and J. S. Levine (eds.). *Case in Slavic*. Columbus, OH: Slavica Publishers, Inc., 86–106.

——1989. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.

——1991. Form and Function in Identifying Cases. In F. Plank (ed.). *Paradigms: The Economy of Inflection*. Berlin: Mouton de Gruyter. 41–55.

——1999. Spatial Cases in Daghestanian Languages. In *Sprachtypologie und Universalienforschung* 52.2. 108–17.

——2005. Alignment of Case Marking. In M. Haspelmath et al. (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 398–405.

—— and M. Polinsky. 1998. The Great Daghestanian case hoax. In A. Siewierska and J. J. Song (eds.). *Case, Typology and Grammar*. Amsterdam/Philadelphia: John Benjamins. 95–114.

—— —— and R. Rajabov. 1998. Tsezian languages. In A. Harris and R. Smets (eds.). *The Languages of the Caucasus*. London: Curzon.

—— G. Stone, and M. S. Polinsky. 1996. *The Russian Language in the Twentieth Century*. Oxford: Oxford University Press.

Cook, K. W. 1988. *A Cognitive Analysis of Grammatical Relations, Case, and Transitivity in Samoan*. PhD dissertation, University of California, San Diego.

Cook, W. A. 1989. *Case Grammar Theory*. Washington: Georgetown University Press.

Corbett, G. G. 1983. *Hierarchies, Targets and Controllers. Agreement Patterns in Slavic*. London: Croom Helm.

—— 2003. Agreement: The Range of the Phenomenon and the Principles of the Surrey Database of Agreement. In *Transactions of the Philological Society* 101. 155–202.

—— 2006. *Agreement*. Cambridge: Cambridge University Press.

Corley, M. and C. Scheepers. 2002. Syntactic Priming in English Sentence Production: Categorical and Latency Evidence from an Internet-based Study. In *Psychonomic Bulletin & Review* 9. 126–31.

Cover, T. M. and J. A. Thomas. 1991. *Elements of Information Theory*. New York: John Wiley & Sons.

Cowper, E. 1988. What is a Subject? Non-nominative Subjects in Icelandic. In J. Blevins and J. Carter (eds.). *Proceedings of NELS*, GLSA, Amherst. 94–108.

Crain, S. 1991. Language Acquisition in the Absence of Experience. *Behavioral and Brain Sciences* 14. 597–611.

—— and J. D. Fodor 1985. How can Grammars Help Parsers? In D. R. Dowty, L. Karttunen, and A. Zwicky (eds.). *Natural Language Parsing. Psychological, Computational and Theoretical Perspectives*. Cambridge: Cambridge University Press. 94–129.

Crass, J. 2005. *Das K'abeena–Deskriptive Grammatik einer hochlandostkuschitischen Sprache*. Cushitic Language Studies, 23. Cologne: Rüdiger Köppe.

Crazzolara, J. P. 1960. *A Study of the Logbara (Ma'di) Language*. Oxford: Oxford University Press.

Creissels, D. 2006a. Encoding the Distinction between Location, Source, and Destination: A Typological Study. In M. Hickmann and S. Robert (eds.). *Space in Languages: Linguistic Systems and Cognitive Categories*. Amsterdam/Philadelphia: John Benjamins. 19–28.

—— 2006b. Cas et postpositions en hongrois. In *Bulletin de la société de linguistique de Paris*, 101. 225–72.

Croft, W. 1988. Agreement vs. Case Marking and Direct Objects. In M. Barlow and C. A. Ferguson (eds.). *Agreement in Natural Language*. Stanford, CA: Centre for the Study of Language and Information. 159–79.

—— 1990. *Typology and Universals*. Cambridge: Cambridge University Press.

—— 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago/London: The University of Chicago Press.

—— 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.

—— 2003. *Typology and Universals*. Cambridge: Cambridge University Press.

Crowley, T. 1998a. *An Erromangan (Sye) Grammar*: Oceanic Linguistics Special Publication, no. 27. Honolulu: University of Hawaii Press.

—— 1998b. *Ura*. Munich/Newcastle: Lincom Europa.

Csató, É. Á. 2000. Syntactic code-copying in Karaim. In Ö. Dahl and M. Koptjevskaja-Tamm (eds.). *The Circum-Baltic Languages: Their Typology and Contacts*. Amsterdam: John Benjamins. 270–83.

Csúcs, S. 1998. Udmurt. In D. Abondolo (ed.). *The Uralic Languages*. London: Routledge, 276–304.

Culicover, P. W. and R. Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Culy, C. 1995. Ambiguity and Case Marking in Donno So (Dogon). In A. Akinalabi (ed.). *Theoretical Approaches to African Linguistics.* Trenton, NJ: Africa World Press. 47–58.

Curnow, T. J. 1997. *A Grammar of Awa Pit (Cuaiquer): An Indigenous Language of Southwestern Colombia*. PhD Thesis. Australian National University.

Cyhun G. A. 1981. *Tipologičeskie problemy balkanoslavjanskogo areala*. Minsk: Navukai tehnika.

Cysouw, M. 2005. Quantitative Methods in Typology. In G. Altmann, R. Köhler, and R. Piotrowski (eds.). *Quantitative Linguistics: An International Handbook*. Berlin: Mouton de Gruyter. 554–78.

Dąbrowska, E. and M. Szczerbiński. 2006. Polish Children's Productivity with Case Marking: The Role of Regularity, Type Frequency, and Phonological Diversity. *Journal of Child Language* 33. 559–97.

Dalrymple, M. 1993. *The Syntax of Anaphoric Binding*. Stanford: CSLI Publications.

—— (ed.). 1999. *Semantics and Syntax in Lexical Functional Grammar*. Cambridge: The MIT Press.

—— 2000. *Lexical Functional Grammar. Syntax and Semantics* Volume 34. New York: Academic Press.

—— and I. Nikoleava. 2006. *Topicality and Nonsubject Marking: Agreement, Casemarking, and Grammatical Function*. Ms., Oxford University.

Daniel, M. A. 2001. Imja susčestvitel'noe. In A. E. Kibrik et al. (eds.). *Bagvalinskij jazyk: Grammatika. Teksty. Slovari.* Moscow: Nasledie. 127–50.

—— 2002. Bagvalinskije toponimy kak problema tipologii chastej rechi. In V. Plungian (ed.). *Issledovanija po teorii grammatiki 2. Grammatikalizacija prostranstvennyx znachenij*. Moscow.

Maisak, T. and S. Merdanova. To appear. *Causatives in Agul.* V. Solougev, and P. Suikhonen (eds.). *Selected Proceedings of the International Symposium on the Typology of Argument Structure and Grammatical Relations in Languages Spoken in Europe and North and Central Asia (LENCA-2). Amsterdam: John Benjamins.*

Davidson, D. 1967. The Logical Form of Action Sentences. In N. Rescher (ed.). *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press. 81–95.

Davidse, K. and B. Lamiroy (eds.). 2002. *The Nominative and Accusative and their Counterparts*. Amsterdam: Benjamins.

Davison, A. 1999. Ergativity: Functional and Formal Issues. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan, and K. Wheatley (eds.). *Functionalism and Formalism in Linguistics, Volume I: General Papers*. Amsterdam: John Benjamins. 177–208.

Dayley, J. P. 1985. *Tzutujil Grammar.* Berkeley: University of California Press.

—— 1989. *Tümpisa (Panamint) Shoshone Grammar*. Berkeley: University of California Press.

DeLancey, S. 1991. Event Construal and Case Role Assignment. In *Berkeley Linguistic Society* 17. 338–53.

—— 2003. Classical Tibetan. In G. Thurgood and R. J. LaPolla (eds.). *The Sino-Tibetan Languages*. London: Routledge. 255–69.

Delbrück, B. 1867. *Ablativ localis instrumentalis im Altindischen*. Berlin: Dümmler.

Dell, G. S. 1986. A Spreading Activation Theory of Retrieval in Language Production. In *Psychological Review* 93. 283–321.

—— F. Chang, and Z. M. Griffin. 1999. Connectionist Models of Language Production: Lexical Access and Grammatical Encoding. In *Cognitive Science* 23. 517–42.

DELSING, L.-O. 1991. Om Genitivens Utveckling i Fornsvenskan [On the Development of the Genitive in Old Swedish]. In S.-G. Malmgren and B. Ralph (eds.). *Studier i Svensk Språkhistoria 2*. Göteborg: Acta Universitatis Gothoburgensis. 12–30.

DENCH, A. C. 1995. *Martuthunira: a Language of the Pilbara Region of Western Australia* (Pacific Linguistics C125). Canberra: Australian National University.

——1998. The Dative Split in Martuthunira. In *Proceedings of the Australian Linguistics Society Meeting 1998*. Available at <http://www.cltr.uq.edu.au/als98/>.

——2006. Case Marking Strategies in Subordinate Clauses in Pilbara Languages: Some Diachronic Speculations. In *Australian Journal of Linguistics* 26. 81–105.

—— and N. R. EVANS 1988. Multiple Case-marking in Australian Languages. In *Australian Journal of Linguistics* 8. 1–47.

DENWOOD, P. 1999. *Tibetan*. Amsterdam: John Benjamins.

DEO, A. and D. SHARMA. 2005. Typological Variation in the Ergative Morphology of Indo-Aryan Languages. To appear in *Linguistic Typology*.

DÉPREZ, V. 1989. *On the Typology of Syntactic Projections and the Nature of Chains: Move-a to the Specifier of Functional Projections*. Ph.D. Dissertation, MIT, Cambridge, MA.

DERDRICK, J. M. and E. H. CASAD. 1999. *Sonora Yaqui Language Structures*. Tucson: University of Arizona.

DICKINS, P. (n.d.). *Ju|'hoan Grammar*. Unpublished ms.

DIK, S. C. 1997. *The Theory of Functional Grammar*. 2nd. rev. ed. Dordrecht: Foris.

DILLER, A. V. N. 1997. Does Thai Permit Detransitivity? In A. S. Abramson (ed.). *Southeast Asian Linguistic Studies in Honour of Vichin Panupong*. Bangkok: Chulalongkorn University Press. 57–78.

DIMMENDAAL, G. J. 1983. *The Turkana Language*. Publications in African Languages and Linguistics 2. Dordrecht and Cinnaminson: Foris Publications.

——2003. Locatives as Core Constituents. In E. Shay and U. Seibert (eds.). *Motion, Direction and Location in Languages*. Amsterdam/Philadelphia: John Benjamins. 91–109.

—— and M. LAST (eds.). 1998. *Surmic Languages and Cultures*. Nilo-Saharan 13 Cologne: Rüdiger Köppe.

DIXON, R. M. W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.

——1979. Ergativity. In *Language* 55. 59–138.

——1980. *The Languages of Australia*. Cambridge: Cambridge University Press.

——1994. *Ergativity*. Cambridge: Cambridge University Press.

——2000. A Typology of Causatives: Form, Syntax and Meaning. In R. M. W. Dixon, and A. Aikhenvald (eds.). *Changing Valency: Case Studies in Transitivity*. Cambridge: Cambridge University Press. 30–83.

——2002. *Australian Languages: Their Nature and Development*. Cambridge: Cambridge University Press.

—— and A. Y. AIKHENVALD. 2000. Introduction. In R. M. W. Dixon and A. Y. Aikhenvald (eds.). *Changing Valency. Case Studies in Transitivity*. Cambridge: Cambridge University Press. 1–29.

DONALDSON, T. 1980. *Ngiyambaa, the Language of the Wangaaybuwan*. Cambridge: Cambridge University Press.

DONEGAN, P. J. and D. STAMPE. 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. 337–53.

Donohue, M. 1999. *A Grammar of Tukang Besi*. Grammar Library Series No. 20. Berlin: Mouton de Gruyter.

—— 2002. Voice in Tukang Besi and the Austronesian Voice System. In F. Wouk and M. Ross (eds.). *The History and Typology of Western Austronesian Voice Systems*. Canberra: Pacific Linguistics 518. 81–99.

—— 2004. Voice Oppositions Without Voice Morphology. In P. Law (ed.). *Proceedings of AFLA 11, ZAS, Berlin 2004*. ZAS Papers in Linguistics Nr. 34–October 2004. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung. 73–88.

—— 2005a. Configurationality in the Languages of New Guinea. In *Australian Journal of Linguistics* 25/2. 237–74.

—— 2005b. Structure is not Syntax: Passive Functions in Tukang Besi. In J. Heinz and D. Ntelitheos (eds.). *Proceedings of the 12th Annual Conference of the Austronesian Formal Linguistics Association*. University of California at Los Angeles Working Papers in Linguistics 12. 73–89. Available online at http://www.linguistics.ucla.edu/faciliti/wpl/issues/wpl12/wpl12.htm

—— 2006. *Verb-initial in Nonverbal Clauses?* Paper presented to the 13th Annual Conference of the Austronesian Formal Linguistics Association. National Tsing Hua University, Hsinchu, Taiwan.

—— 2007. Stative–Active Systems: What's What and What's Not. In M. Donohue and S. Wichmann (eds.). *Typology of Languages with Semantic Alignment*. Oxford: Oxford University Press.

—— and L. Brown. 1999. Ergativity: Some Additions from Indonesia. In *Australian Journal of Linguistics* 19. 57–76.

—— and S. Wichmann (eds.). 2007. *Typology of Languages with Semantic Alignment*. Oxford: Oxford University Press.

Dowty, D. R. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

—— 1982. Grammatical Relations and Montague Grammar. In G. K. Pullum and P. Jacobson (eds.). *The Nature of Syntactic Representation*. Dordrecht: Reidel. 79–130.

—— 1989. On the Semantic Content of the Notion of 'Thematic Role'. In G. Chierchia, B. H. Partee, and R. Turner (eds.). *Properties, Types and Meanings*, 2: Semantic Issues. Dordrecht: Kluwer. 69–129.

—— 1991. Thematic Proto-Roles and Argument Selection. In *Language* 67(3). 547–619.

Drai, D. and Y. Grodzinsky. 2006. A New Empirical Angle on the Variability Debate: Quantitative Neurosyntactic Analyses of a Large Data Set from Broca's Aphasia. In *Brain and Language* 76(2). 117–28.

Dressler, W. U. (ed.). 1997. *Studies in Pre- and Protomorphology*. Vienna: Verlag der Österreicheschen Akademie der Wissenschaften.

Dressler, W. U. 2005. Morphological Typology and First Language Acquisition: Some Mutual Challenges. In G. Booij, E. Guevara, A. Ralli, S. Sgroi, and S. Scalise (eds.). *MMM4 Proceedings: Morphology and Linguistic Typology. On-line Proceedings of the Fourth Mediterranean Morphology Meeting (MMM4) Catania, 21–23 September 2003*. http://morbo.lingue.unibo.it/mmm/proc-mmm4.php

—— and A. Karpf. 1995. The Theoretical Relevance of Pre- and Protomorphology in Language Acquisition. *Yearbook of Morphology* 1994. 99–124.

Au: Please check updation against Ref. in Ch. 24

van Driem, G. 1987. *A Grammar of Limbu.* Berlin: Mouton de Gruyter.
—— 1993. *A Grammar of Dumi*. Berlin/New York: Mouton de Gruyter.
Drossard, W. 1991. Verbklassen. In H. Seiler and W. Premper (eds.). *Partizipation: das sprachliche Erfassen von Sachverhalten*. Tübingen: Narr. 150–82.
Dryer, M. 1986. Primary Objects, Secondary Objects, and Antidative. In *Language* 62. 808–45.
—— 1989. Large Linguistic Areas and Language Sampling. In *Studies in Language* 13. 257–92.
—— 2002. Case Distributions, Rich Agreement and Word Order Type. In *Theoretical Linguistics* 28. 151–7.
—— 2005a. Prefixing Versus Suffixing in Inflectional Morphology. In M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 110–13.
—— 2005b. Position of Case Affixes. In M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 210–13.
—— 2005c. The Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase. In M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 386–9.
—— 2007. Clause Types. In: T. Shopen (ed.). *Language Typology and Syntactic Description*, 2nd edition. Cambridge: Cambridge University Press.
Dubnova, Y. Z. 1984. *The Rwanda Language*. Moscow: Nauka.
—— Du Bois, J. A. 1985. Competing Motivations. In J. Haiman (ed.). *Iconicity in Syntax*. Amsterdam: Benjamins. 343–66.
Du Bois, J. W. 1987. The Discourse Basis of Ergativity. In *Language* 63. 805–55.
Dunn, J. A. 1979. *A Reference Grammar for the Coast Tsimshian Language*. Canadian Ethnology Service, Mercury Series Paper No. 55. Ottawa: National Museum of Man.
Durie, M. 1995. Towards an Understanding of Linguistic Evolution and the Notion 'X has a function Y'. In W. Abraham, T. Givón, and S. A. Thompson (eds.). *Discourse Grammar and Typology (Papers in honour of John W. M. Verhaar)*. Amsterdam: Benjamins. 275–309.
Džaukjan (Djahukian), G. B. 1982. *Sravnitel'naja grammatika armjanskogo jazyka* [A comparative grammar of Armenian]. Erevan: AN ArmSSR.
Ebeling, C. L. 1966. Review of Chikobava and Cercvadze's 'The Grammar of Literary Avar'. In *Studia Caucasica* 2. 58–100.
Ebert, K. 1997a. *Camling*. Munich: Lincom.
—— 1997b. *A Grammar of Athpare*. Munich: Lincom Europa.
Édel'man, D. I. 1966. *Jazguljamskij jazyk*. Moscow: Nauka.
—— 1983. *The Dardic and Nuristani Languages*. Moscow: Nauka.
Edzard, D. O. 2003. *Sumerian Grammar*. Leiden: Brill.
Eisenbeiss, S. 1994a. Auxiliaries and the Acquisition of the Passive. In *The Proceedings of the 25th Annual Child Language Research Forum*. Stanford, CA: CSLI Publications. 235–42.
—— 1994b. Kasus und Wortstellungsvariation im deutschen Mittelfeld; theoretische Überlegungen und Untersuchungen zum Erstspracherwarb. In B. Haftka (ed.). *Was determiniert Wortsellungsvariationen?* Opladen: Westdeutscher Verlag. 277–98.
—— 2000. The Acquisition of the Determiner Phrase in German Child Language. In M.-A. Friedemann, and L. Rizzi (eds.). *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics.* London: Longman. 26–62.

——2003. *Merkmalsgesteuerter Grammatikerwerb*. Dissertation, University of Düsseldorf. (An electronic copy can be downloaded from: http://deposit.ddb.de/cgi-bin/dokserv?idn=97646330x&dok_var=d1&dok_ext=pdf&filename=97646330x.pdf)

——S. Bartke, and H. Clahsen. 2006. Structural and Lexical Case in Child German: Evidence from Language-Impaired and Typically-Developing Children. In *Language Acquisition* 13. 3–32.

Elfenbein, J. 1990. *An Anthology of Classical and Modern Balochi Literature*. 2 vols. Wiesbaden: Harrassowitz Verlag.

Emergence. 2005. Emergence of Nominal and Verbal Morphology from a Typological Perspective. *Symposium. Xth International Congress for the Study of Child Language, July 2005. Abstracts*. Berlin.

Emonds, J. 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.

——1987. The Invisible Category Principle. In *Linguistic Inquiry* 18. 613–32.

Enç, M. 1991. The Semantics of Specificity. In *Linguistic Inquiry* 22. 1–25.

Enfield, N. J. 2002. Combinatoric Properties of Natural Semantic Metalanguage Expressions in Lao. In C. Goddard and A. Wierzbicka (eds.). *Meaning and Universal Grammar: Theory and Empirical Findings*. vol. 2. Amsterdam: John Benjamins. 145–256.

——2005. Areal Linguistics and Mainland Southeast Asia. In *Annual Review of Anthropology* 34. 181–206.

——(in press, a) *A Grammar of Lao*. Berlin: Mouton de Gryuter.

——(in press, b). Encoding Three-Participant Events in the Lao Clause. In *Linguistics*.

Engelberg, S. 2000. *Verben, Ereignisse und das Lexikon*. Tübingen: Niemeyer.

England, N. C. 1983. *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press.

——1988. Mam Voice. In M. Shibatani (ed.). *Passive and Voice*. Amsterdam: John Benjamins. 525–46.

Epstein, S. D. and T. D. Seely. 2006. *Derivations in Minimalism*. Cambridge: Cambridge University Press.

Erelt, M., T. Erelt, and K. Ross. 2000. *Eesti keele käsiraamat*. Tallinn: Eesti Keele Sihtasutus.

Erteschik-Shir, N., and T. Rapoport (eds.). 2005. *The Syntax of Aspect. Deriving Thematic and Aspectual Interpretation*. Oxford: Oxford University Press.

Evans, N. 1985. *Kayardild: The Language of the Bentinck Islanders of North-west Queensland*. PhD thesis, Australian National University.

——1995a. *A Grammar of Kayardild. With comparative notes on Tangkic*. Berlin: Mouton de Gruyter.

——1995b. Multiple Case in Kayardild: Anti-iconic Suffix Ordering and the Diachronic Filter. In F. Plank (ed.). *Double Case. Agreement by Suffixaufnahme*. New York, Oxford: Oxford University Press. 396–428.

——and R. Nordlinger (2004) *Extreme Morphological shift: Verbal Case in Kayardild*. Talk presented at the LFG04 meeting, Christchurch, New Zealand, July 2004. Unpublished ms., University of Melbourne.

Eythórsson, T. 2002. Changes in Subject Case-Marking in Icelandic. In D. Lightfoot (ed.). *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press. 196–212.

Fajzov, M. 1966. *Jazyk rushancev Sovetskogo Pamira*. Dushanbe: Akademija nauk Tadzhikskoj SSSR.

Falk, C. 1997. *Fornsvenska Upplevarverb* [Experiencer verbs in Old Swedish]. Lund: Lund University Press.

Falk, Y. N. 1997. *Case Typology and Case Theory*. Ms., Hebrew University, Jerusalem.

Faust, N. 1973. *Lecciones para el aprendizaje del idioma Shipibo-Conibo* [Lessons in Shipibo-Conibo]. *Documento de Trabajo*, 1. Yarinacocha: Instituto Lingüístico de Verano.

Feldman, H. 1986. *A Grammar of Awtuw*. Canberra: Australian National University.

Felix, R. 2005. *A Grammar of River Walihío*. PhD Thesis. Rice University.

Felix, S. W. 1987. *Cognition and Language Growth*. Dordrecht: Foris.

——1992. Language Acquisition as Maturational Process. In J. Weissenborn, H. Goodluck, and T. Roeper (eds.). *Theoretical Issues in Language Acquisition: Continuity and Change in Development*. Hillsdale, NJ: Erlbaum. 25–51.

Feoktistov, A. P. 1966. Erzjanskij jazyk. In *Jazyki Narodov SSSR*, vol. 3. Moscow: Nauka, 177–98.

Ferreira, F. 1994. The Creation of Prosody during Sentence Production. In *Psychological Review* 100. 233–53.

——2000. Syntax in Language Production: An Approach using Tree-adjoining Grammars. In L. Wheeldon (ed.). *Aspects of Language Production*. Cambridge, MA: MIT Press. 291–330.

——and B. Swets. 2002. How Incremental is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. In *Journal of Memory and Language* 46. 57–84.

Ferreira, V. S. 1996. Is it Better to Give than to Donate? The Consequences of Syntactic Flexibility in Language Production. In *Journal of Memory and Language* 35. 724–55.

——and G. S. Dell. 2000. Effects of Ambiguity and Lexical Availability on Syntactic and Lexical Production. In *Cognitive Psychology* 40. 296–340.

Fife, J. and G. King. 1998. Celtic (Indo-European). In A. Spencer and A. Zwicky (eds.). *Handbook of Morphology*. Oxford: Blackwell Publishers. 477–99.

Filliozat, P.-S. 1988. *Grammaire Sanskrite Pâninéenne*. Paris: Picard.

Fillmore, C. J. 1965. Toward a Modern Theory of Case. *Project on LinguisticAnalysis, Ohio State University* 13. 1–24. [Reprinted in: D. A. Reibel and S. A. Schane (eds.). *Modern Studies in English*. Englewood Cliffs, New Jersey: Prentice-Hall, 1969. 361–75.]

——1966. A Proposal concerning English Prepositions. In *Georgetown University Monograph Series on Languages and Linguistics* 19. 19–33.

——1968. The Case for Case. In E. Bach and R. T. Harms (eds.). *Universals of Linguistic Theory*. New York: Holt, Rinehart and Winston. 1–90.

——(1972). Subjects, Speakers, and Roles. In D. Davidson and G. Harman (eds.). *Semantics of Natural Language*. Dordrecht: Reidel. 1–24.

——1977. The Case for Case Reopened. In P. Cole and J. Sadock (eds.). *Syntax and Semantics: Grammatical Relations*, Volume 8. New York: Academic Press. 59–91.

——P. Kay, and M. C. O'Connor. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of *let alone*. In *Language* 64.3. 501–38.

Fleming, H. C. 1976. Omotic Overview. In M. Bender (ed.). *The Non-Semitic Languages of Ethiopia*. East Lansing : African Studies Center, Michigan State University. 299–323.

Fodor, J. D. 1978. Parsing Strategies and Constraints on Transformations. *Linguistic Inquiry* 9.

Foley, W. A. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.

——1991. *The Yimas Language of New Guinea*. Stanford, CA: Stanford University Press.

——1998. *Symmetrical Voice Systems and Precategoriality in Philippine Languages.* Paper presented at the LFG98 Conference, Brisbane.
——2000. The Languages of New Guinea. In *Annual Review of Anthropology* 29. 357–404.
——and R. D. Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
Fong, V. 1997. *The Order of Things: What Directional Locatives Denote*. PhD dissertation, Stanford University.
Foris, D. P. 2000. *A Grammar of Sochiapan Chinantec*. Studies in Chinantec Languages 6. Arlington: SIL International and The University of Texas at Arlington.
Forssman, B. 2001. *Lettische Grammatik*. Münchener Studien zur Sprachwissenschaft 20. Dettelbach: Röll.
Fortescue, M. 1984. *West Greenlandic*. Croom Helm Descriptive Grammars. London/Sydney Dover: Croom Helm.
Foulet, L. 1970. *Petite syntaxe de l'ancien français*. 3rd rev. edn. Paris: Champion.
Frajzyngier, Z. 2002. *A Grammar of Hdi*. Berlin/New York: Mouton de Gruyter.
Franck, J., G. Lassi, U. H. Frauenfelder, and L. Rizzi. 2006. Agreement and Movement: A Syntactic Analysis of Attraction. In *Cognition* 101. 173–216.
Frank, P. S. 1990. *Ika Syntax*. Studies in the languages of Colombia 1. Arlington, TX: Summer Institute of Linguistics.
Frank, R. 1992. *Syntactic Locality and Tree Adjoining Grammar: Grammatical, Acquisition and Processing Perspectives*. PhD thesis, University of Pennsylvania.
Fraurud, K. 2001. Possessives with Extensive Use: A Source of Definite Articles? In I. Baron, M. Herslund, and F. Sørensen (eds.). *Dimensions of Possession*. Amsterdam: John Benjamins. 243–67.
Frazier, L. 1987. Syntactic Processing: Evidence from Dutch. In *Natural Language and Linguistic Theory* 5. 519–59.
——and K. Rayner. 1982. Making and Correcting Errors During Sentence Comprehension: Eye Movements in the Analysis of Structurally Ambiguous Sentences. In *Cognitive Psychology* 14. 178–210.
Freidin, R. and R. A. Sprouse. 1991. Lexical Case Phenomena. In R. Freidin (ed.). *Principles and Parameters in Comparative Grammar*. Cambridge, MA: MIT Press. 392–416.
Frey, J. 2003. *Ellipsis and wa-Marking in Japanese Conversation*. New York: Routledge.
Frey, W. 2001 About the Whereabouts of Indefinites. In *Theoretical Linguistics* 27. 137–61.
Friederici, A. D. and A. Mecklinger. 1996. Syntactic Parsing as Revealed by Brain Responses: First-Pass and Second-Pass Parsing Processes. In *Journal of Psycholinguistic Research* 25. 157–76.
Friedmann, N. 2001. Agrammatism and the Psychological Reality of the Syntactic Tree. In *Journal of Psycholinguistic Research* 30. 71–90.
——2006. Speech Production in Broca's Agrammatic Aphasia: Syntactic Tree Pruning. In Y. Grodzinsky and K. Amunts (eds.). *Broca's Region.* New York: Oxford University Press. 63–82.
——and L. Shapiro. 2003. Agrammatic Comprehension of Simple Active Sentences with Moved Constituents: Hebrew OSV and OVS Structures. In *Journal of Speech, Language and Hearing Research* 46. 288–97.
Fromkin, V. A. 1971. The Non-anomalous Nature of Anomalous Utterances. In *Language* 47. 27–52.
——1973. *Speech Errors as Linguistic Evidence*. The Hague: Mouton.

Fukui, N. and M. Speas. 1986. Specifiers and Projection. In N. Fukui, T. R. Rapoport, and E. Sagey (eds.). *Papers in Theoretical Linguistics*. MIT Working Papers in Linguistics 8. 128–72.

Gair, J. W. 1990. Subjects, Case and INFL in Sinhala. In M. Verma and K. P. Mohanan (eds.). *Experiencer Subjects in South Asian Languages*. Stanford: CSLI. 13–42.

Galand, L. 1964. L'énoncé verbal en berbère. In *Cahiers Ferdinand de Saussure* 21. 33–53.

Ganenkov, D. S. 2005. *Kontaktnye lokalizacii v naxsko-dagestanskix jazykax i ix tipologičeskie paralleli*. PhD thesis. State University of Moscow.

—— (2006). Experiencer Coding in Nakh-Daghestanian. In L. Kulikov et al. (eds.). *Case, Valency and Transitivity*. Studies in Language Companion Series. Amsterdam: John Benjamins.

—— T. Maisak, and S. Merdanova. 2006. *Non-Canonical Agent Marking in Agul*. Ms., Moscow State University.

Garnsey, S., M. Tanenhaus, and R. Chapman. 1989. Evoked Potentials and the Study of Sentence Comprehension. In *Journal of Psycholinguistic Research* 18. 51–60.

Garrett, A. 1990. The Origin of NP Split Ergativity. In *Language* 66/2. 261–96.

Garrett, M. F. 1975. The Analysis of Sentence Production. In G. H. Bower (ed.). *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 9. New York: Academic Press. 133–77.

—— 1980. Levels of Processing in Sentence Production. In B. Butterworth (ed.). *Language Production*, Vol. 1: Speech and Talk. London: Academic Press. 177–220.

Gary, J. O. and S. Gamal-Eldin. 1982. *Cairene Egyptian Colloquial Arabic*. Amsterdam: North-Holland.

Gasser, M. 1983. Topic Continuity in Written Amharic Narrative. In T. Givón (ed.). *Topic Continuity in Discourse: A Quantitative Cross-language Study*. Amsterdam/Philadelphia: John Benjamins. 95–139.

Gauto Bejarano, M. G. 1990. *Introducción a la gramática elemental del guaraní*. Asunción: El Foro.

Gazdar, G., E. Klein, G. Pullum, and I. Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

van Geenhoven, V. 1996. *Semantic Incorporation and Indefinite Descriptions: Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. PhD dissertation, University of Tübingen.

van Gelderen, V. 2003. *Scrambling Unscrambled*. PhD dissertation, Leiden University.

Genetti, C. 1991. From Postposition to Subordinator in Newari. In E. C. Traugott and B. Heine (eds.). *Approaches to Grammaticalization*, Volume 2. Amsterdam/Philadelphia: Benjamins. 227–55.

Gensler, O. 2002. Order of Double Objects. In *Linguistic Typology* 6.3. 35–79.

Georg, S. 1996. *Marphatan Thakali: Untersuchungen zur Sprache des Dorfes Marpha im Oberen Kāli-Gandāki-Tal/Nepal*. Munich: Lincom Europa.

Geschwind, N. 1965. Disconnexion Syndromes in Animals and Man. In *Brain* 88. 585–644.

van Gestel, F., J. N. Twilhaar, T. Rinkel, and F. Weerman. 1992. *Oude zinnen; grammaticale analyse van het Nederlands tussen 1200–1700*. Leiden, Martinus Nijhoff.

Gil, D. 2005a. Isolating-Monocategorial-Associational Language. In H. Cohen and C. Lefebvre (eds.). *Categorization in Cognitive Science*. Amsterdam: Elsevier.

——2005b. Word Order without Syntactic Categories: How Riau Indonesian Does It. In A. Carnie, S. D. Collberg, and H. Harley (eds.). *Verb First: Papers on the Syntax of Verb Initial Languages*. Amsterdam: John Benjamins.

Gildersleeve, B. L. and G. Lodge. 1968. *Latin Grammar* (3rd edition). London: Macmillan.

Givón, T. 1975. Serial Verbs and Syntactic Change: Niger-Congo. In C. Li (ed.). *Word Order and Word Order Change*. Austin/London: University of Texas Press. 47–111.

——1979. *On Understanding Grammar*. New York: Academic Press.

——1980. *Ute Reference Grammar*. Ignacio, CO: Ute Press.

——1984. *Syntax: A Functional-Typological Introduction*, Vol. I. Amsterdam: Benjamins.

——1985. Ergative Morphology and Transitivity Gradients in Newari. In F. Plank (ed.). *Relational Typology*. Berlin: Mouton. 89–107.

——1988. The Pragmatics of Word Order: Predictability, Importance and Attention. In M. Hammond et al. (eds.). *Studies in Syntactic Typology*. Amsterdam: John Benjamins. 243–84.

——1995. *Functionalism and Grammar*. Amsterdam/Philadelphia: John Benjamins.

——2001. *Syntax. A Functional-Typological Introduction*, vol. 2. Amsterdom: John Benjamins.

——2006. *Multiple Routes to Clause Union: the Diachrony of Syntactic Complexity.* Ms. Eugene, University of Oregon.

Gladkova, A. Forthcoming. *Russian Emotions, Attitudes and Values: Selected Topics in Cultural Semantics*. PhD Thesis. Canberra: Australian National University.

Goddard, C. 1982. Case Systems and Case Marking in Australian Languages: A New Interpretation. In *Australian Journal of Linguistics* 2. 167–96.

——1998. *Semantic Analysis: A Practical Introduction*. Oxford: Oxford University Press.

——2003. Dynamic -ter in Malay (Bahasa Melayu): A study in grammatical polysemy. In *Studies in Language* 27(2). 287–322.

——and A. Wierzbicka. 1994. Introducing Lexical Primitives. In C. Goddard and A. Wierzbicka (eds.). *Semantic and Lexical Universals*. Amsterdam: John Benjamins. 31–54.

—— ——(eds.). 2002a. *Meaning and Universal Grammar: Theory and Empirical Findings*. 2 volumes. Amsterdam: John Benjamins.

—— ——2002b. Semantic Primes and Universal Grammar. In C. Goddard and A. Wierzbicka (eds.). *Meaning and Universal Grammar: Theory and Empirical Findings*, vol. I. Amsterdam: John Benjamins. 41–86.

—— ——In press. Semantic Primes and Cultural Scripts in Language Learning and Intercultural Communication. In F. Sharifian and G. Palmer (eds.). *Applied Cultural Linguistics*. Amsterdam: John Benjamins.

Godel, R. 1975. *An Introduction to the Study of Classical Armenian*. Wiesbaden: Reichert.

Goldberg, A. E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goodglass, H. and E. Kaplan. 1983. *The Assessment of Aphasia and Related Disorders* (2nd edition). Philadelphia: Lea & Febiger.

Gordon, L. 1986. *Maricopa Morphology and Syntax*. University of California Publications in Linguistics, 108. Berkeley: University of California Press.

Gorelova, L. M. 2002. *Manchu Grammar*. Leiden and Boston: Brill.

GRAMATIKA 1982–1983. *Gramatika na săvremennija bălgarski knižoven ezik: v tri toma*. S. Stojanov (eds.). Sofia: Balgarskata akademija na naukite.

GRASHCHENKOV, P. V. 2006. *Sintaksis i tipologija genitivnoj gruppy*. Dissertation, Moscow State University.

GREENBERG, J. H. 1959. The Origin of the Maasai Passive. In *Africa* 29, 2. 171–6.

—— 1963a. *The Languages of Africa*. The Hague: Mouton.

—— 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.

—— 1966. Some Universals of Grammar, with Particular Reference to the Order of Meaningful Elements. In J. H. Greenberg (ed.). *Universals of Language*. Cambridge, MA: MIT. 73–113.

GREENFIELD, S. 2001. *Brain Story*. Baarn: Bosch & Keuning.

GRÉGOIRE, C. 1975. *Les locatifs en Bantou*. Tervuren: Musée Royal de l'Afrique Centrale.

GREGORES, E. and J. A. SUÁREZ. 1967. *A Description of Colloquial Guaraní*. The Hague: Mouton.

GRENAND, F. 1980. *La langue wayãpi (Guyane française). Phonologie et grammaire.* Paris: SELAF.

GREWENDORF, G. and J. SABEL. 1999. Scrambling in German and Japanese: Adjunction versus Multiple Specifiers. In *Natural Language and Linguistic Theory* 17/1. 1–86.

GRIES, S. 2005. Syntactic Priming: A Corpus-based Approach. *Journal of Psycholinguistic Research* 34. 365–99.

GRIMM, S. 2005. *The Lattice of Case and Agentivity*. MA Thesis, University of Amsterdam.

GRIMSHAW, J. B. 1990. *Argument Structure*. Cambridge, MA: MIT Press.

GRJUNBERG, A. L. 1963. *Jazyk Severoazerbajdžanskix Tatov*. Moscow: Izdatel'stvo Akademii Nauk SSSR.

—— 1980. *Jazyk kati*. [The Kati Language.] Moscow: Vostočnaya Literatura.

—— and D. I. ÈDEL'MAN. 1987. Afganskij Jazyk. In: *Osnovy iranskogo jazykoznanija*, vol. 4: *Novoiranskie jazyki: vostočnaja gruppa*. Moscow: Nauka. 6–154.

GRODZINSKY, Y. 1990. *Theoretical Perspectives on Language Deficits*. Cambridge, MA: MIT Press.

—— 2000. The Neurology of Syntax: Language Use without Broca's Area. In *Behavioral and Brain Sciences* 23. 1–71.

—— and L. FINKEL. 1998. The Neurology of Empty Categories: Aphasics' Failure to Detect Ungrammaticality. In *Journal of Cognitive Neuroscience* 10. 281–92.

—— K. WEXLER, Y. C. CHIEN, S. MARAKOVITZ, and J. SOLOMON. 1993. The Breakdown of Binding Relations. In *Brain and Language* 45. 396–422.

DE GROOT, A. W. 1939. Les oppositions dans les systèmes de la syntaxe et des cas. In *Mélanges de linguistique offerts à Charles Bally*. Geneva: Georg. 107–27.

GRUBER, J. S. 1965. *Studies in Lexical Relations*. Doctoral dissertation, MIT. [Published as *Lexical Structures in Syntax and Semantics*, North-Holland, Amsterdam, 1976.]

GRUBER-MILLER, A. M. 1990. Loss of Nominal Case Endings in the Modern Arabic Sedentary Dialects: Evidence from Southern Palestinian Christian Middle Arabic Texts. In M. Eid (ed.). *Perspectives on Arabic Linguistics I: Papers from the First Annual Symposium on Arabic Linguistics*. Amsterdam: Benjamins. 235–53.

GRÜNTHAL, R. 2003. *Finnic Adpositions and Cases in Change*. Helsinki: Société Finno-Ougrienne.

Gruzdeva, E. 1998. *Nivkh*. Lincom/Materials, 111. Munich: Lincom.

Guerrero, L. and R. D. Van Valin, Jr. 2004. Yaqui and the Analysis of Primary Object Languages. In *International Journal of American Linguistics* 70. 290–319.

Guilfoyle, E., H. Hung, and L. Travis. 1992. Spec of IP and Spec of VP: Two Subjects in Austronesian Languages. In *Natural Language and Linguistic Theory* 10. 375–414.

Guillaume, A. 2004. A *Grammar of Cavineña, an Amazonian Language of Northern Bolivia*. PhD Thesis, Research Centre for Linguistic Typology. La Trobe University (Melbourne, Australia).

Gukasian, V. 1974. *Udinsko-azerbajdžansko-russkij slovar'*. Baku: Elm.

Gumperz, J. J. and R. Wilson. 1971. Convergence and Creolization: A Case from the Indo-Aryan/Dravidian Border in India. In D. Hymes (ed.). *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press. 151–67.

Gvozdanović, J. 1997. Tracing the Origin of a Change. In J. Gvozdanović (ed.). *Language Change and Functional Explanation*. Berlin: Mouton de Gruyter. 71–87.

Gvozdev A. N. 1961. *Voprosy izučenija detskoj reći* [Questions of the study of child language]. Moscow: Prosveščenie.

Hagège, C. 1993. *The Language Builder. An Essay on the Human Structure in Linguistic Morphogenesis*. Amsterdam/Phildelphia: John Benjamins.

Hagiwara, H. and D. Caplan. 1990. Syntactic Comprehension in Japanese Aphasics: Effects of Category and Thematic Role Order. In *Brain and language* 38. 159–70.

Hagman, R. S. 1973. *Nama Hottentot Grammar*. Columbia University.

Hagoort, P., C. M. Brown, and T. Y. Swaab. 1996. Lexical–semantic Event-Related Potential Effects in Patients with Left Hemisphere Lesions and Aphasia, and Patients with Right Hemisphere Lesions without Aphasia. In *Brain* 119. 627–49.

Haider, H. 1993. *Deutsche Syntax–Generativ. Vorstudien zu einer Projektiven Theorie der Grammatik*. Tübingen: Narr.

—— 2006. Mittelfeld Phenomena. In M. Everaert and H. van Riemsdijk (eds.). *The Blackwell Companion to Syntax*. Oxford: Blackwell.

Haider, H. and I. Rosengren. 1998. *Scrambling*. Sprache und Pragmatik 49. Lund.

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.

—— (forthcoming). *Alignment Shift in Iranian languages*. Berlin: Mouton.

Haile, G. 1970. The Suffix Pronouns in Amharic. In C.-W. Kim and H. Stahlke (eds.). *Papers in African Linguistics*. Edmonton: Linguistic Research, Inc. 101–11.

Haiman, J. (ed.). 1985. *Iconicity in Syntax*. Amsterdam.

Hakulinen, L. 1961. *The Structure and Development of the Finnish Language*. Bloomington: Indiana University; The Hague: Mouton.

Hale, K. 1973. Person Marking in Walbiri. In S. Anderson and P. Kiparsky (eds.). *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Wilson. 308–44.

—— 1981. Preliminary Remarks on the Grammar of Part–Whole Relations in Warlpiri. In J. Hollyman and A. Pawley (eds.). *Studies in Pacific Languages and Cultures*. Auckland: Linguistic Society of New Zealand. 333–44.

—— 2002. Eccentric Agreement. In B. Fernández and P. Albizu (eds.). *Kasu eta Komunztaduraren gainean* [On Case and Agreement]. Euskal Herriko Unibetsitatea, Vitoria-Gasteiz. 15–48.

978–0–19–920647–6 59-Spencer-Ref OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 848 of 894 May 16, 2008 19:1

Hale, K. and S. J. Keyser. 1992. The Syntactic Character of Thematic Structure. In I. M. Roca (ed.). *Thematic Structure. Its Role in Grammar*. Berlin: Foris. 107–44.

——— 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In K. Hale and S. J. Keyser (eds.). *The View from Building 20*. Cambridge: The MIT Press. 53–109.

——— 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.

Hall, C. J. 1984. Integrating Diachronic and Processing Principles in Explaining the Suffixing Preference. In J. A. Hawkins (ed.). *Explaining Language Universals*. Oxford: Blackwell. 321–49.

Halle, M. and A. Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In K. Hale and S. J. Keyser (eds.). *The View from Building 20. Essays Presented to Silvain Bromberger on his 50th Birthday*. Cambridge, MA: MIT Press. 111–76.

Halvorsen, P.-K. and R. M. Kaplan. 1988. Projections and Semantic Description in Lexical-Functional Grammar. In *Proceedings of the International Conference on Fifth Generation Computer Technology*. 1116–1122. [Reprinted 1995 in M. Dalrymple et al. (eds.). *Formal Issues in Lexical-Functional Grammar*.]

Han, J. 1999. *On Grammatical Coding of Information Structure in Korean: A Role and Reference Grammar Account*. Unpublished PhD Dissertation, University at Buffalo. Available on RRG web site: http://linguistics.buffalo.edu/research/rrg.html.

Hansson, G. Ó. (2007). Productive Syncretism in Saami Inflectional Morphology. In I. Toivonen and D. Nelson (eds.). *Saami Linguistics*. Amsterdam/Philadelphia: John Benjamins.

Harris, A. 1981. *Georgian Syntax: A Study in Relational Grammar*. Cambridge: Cambridge University Press.

——— 1984. Inversion as a rule of Universal Grammar: Georgian evidence. In D. M. Perlmutter and C. G. Rosen (eds.). *Studies in Relational Grammar 2*. Chicago/London: The University of Chicago Press. 259–91.

——— and L. Campbell. 1995. *Historical Syntax in Cross-linguistic Perspective*. Cambridge Studies in Linguistics 74. Cambridge: Cambridge University Press.

Hartsuiker, R. J. and C. Westenberg. 2000. Word Order Priming in Written and Spoken Sentence Production. In *Cognition* 75. B27–B39.

——— H. Kolk, and W. J. Huinck. 1999a. Agrammatic Production of Subject-Verb Agreement: The Effect of Conceptual Number. In *Brain and Language* 69(2). 119–60.

——— ——— and P. Huiskamp. 1999b. Priming Word Order in Sentence Production. In *Quarterly Journal of Experimental Psychology* 52A. 129–47.

Harvey, M. 2001. *A Grammar of Limilngan: A Language of the Mary River Region, Northern Territory, Australia*. Canberra: Pacific Linguistics.

Hasada, R. 1997. Conditionals and Counterfactuals in Japanese. In *Language Sciences* 19(3). 277–88.

Haspelmath, M. 1993a. *A Grammar of Lezgian*. Berlin: Mouton de Gruyter.

——— 1993b. More on the Typology of Inchoative/Causative Verb Alternations. In B. Comrie and M. Polinsky (eds.). *Causation and Transitivity*. Amsterdam: John Benjamins.

——— 1999a. External Possession in a European Areal Perspective. In: D. L. Payne and I. Barshi (eds.). *External Possession*. Typological Studies in Language 39. Amsterdam: Benjamins. 109–35.

——1999b. Explaining Article-Possessor Complementarity: Economic Motivation in Noun Phrase Syntax. In *Language* 75. 227–43.
——2001. The European Linguistic Area: Standard Average European. In M. Haspelmath et al. (eds.). *Language Typology and Language Universals: An International Handbook*. Berlin: Mouton de Gruyter. 1492–1510.
——2002. *Understanding Morphology*. London: Arnold.
——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: Lawrence Erlbaum. 211–43.
——2004. Explaining the Ditransitive Person–Role Constraint: A Usage Based Account. *Constructions* 2/2004 (free online journal, University of Düsseldorf). 49 pp.
——2005a. Argument Marking in Ditransitive Alignment Types. In *Linguistic Disovery* 3. 1–21. http://linguistic-discovery.dartmouth.edu.
——2005b. Ditransitive Constructions: the Verb Give. In M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 426–9.
——2006. Against Markedness (and What to Replace it With). In *Journal of Linguistics* 41(2). 1–46.
——and S. Michaelis. (forthcoming). Leipzig fourmille de typologues – Genitive Objects in Comparison. In G. G. Corbett and M. Noonan (eds.). *Case and Grammatical Relations: Papers in Honour of Bernard Comrie*. Amsterdam: John Benjamins.
——and T. Müller-Bardey. 2004. Valency Change. In G. Booij, C. Lehmann, and J. Mugdan (eds.). *Morphology: A Handbook on Inflection and Word Formation*. Vol. 2. Berlin: de Gruyter. 1130–45.
——M. S. Dryer, D. Gil, and B. Comrie. 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.
Hasselblatt, C. 1992. *Grammatisches Wörterbuch des Estnischen*. Wiesbaden: Harrassowitz.
Haude, K. 2006. *A Grammar of Movima*. PhD Thesis, Radboud University, Nijmegen.
Hawkins, J. A. (ed.). 1984. *Explaining Language Universals*. Oxford: Blackwell.
——1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
——2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
——and A. Cutler. 1988. Psycholinguistic Factors in Morphological Asymmetry. In J. A. Hawkins (ed.). *Explaining Language Universals*. Oxford: Blackwell. 280–317.
Hawkins, R. E. 1998. Wai Wai. In D. C. Derbyshire and G. K. Pullum (eds.). *Handbook of Amazonian Languages*, vol. 4. Berlin: Mouton de Gruyter. 25–224.
——2001. *Second Language Syntax: a Generative Introduction*. Oxford: Blackwell.
Hayes, J. 1997. *Sumerian*. Lincom/Materials, 68. Munich: Lincom.
Hayward, R. J. 1979. Bayso Revisited; Some Preliminary Linguistic Observations II. In *Bulletin of the School of Oriental and African Studies (BSOAS)* 42. 101–42.
——1984. *The Arbore Language: A First Investigation*. (Including a Vocabulary.) Cushitic Language Studies 2. Hamburg: Helmut Buske.
——1988. Is There a Language with an Indefinite Nominative–Burji. In T. Beyene (ed.). *Proceedings of the Eighth International Conference of Ethiopian Studies*. University of Addis Ababa. 679–91.
——1990. Notes on Zayse Language. In R. Hayward (ed.). *Omotic Language Studies*. London: School of Oriental and African Studies. 210–355.

HAYWARD, R. J. and Y. TSUGE. 1998. Concerning Case in Omotic. In *Afrika und Übersee,* 81. 21–38.

HEATH, J. 2005. *A Grammar of Tamashek (Tuareg of Mali)*. Berlin/New York: Mouton de Gruyter.

HEESCHEN, C. 1980. Strategies of Decoding Actor-Object-Relations by Aphasic Patients. In *Cortex* 16. 5–19.

HEINE, B. 1976. *The Kuliak Languages of Eastern Uganda*. Nairobi: East African Publishing House.

——1983. *The Ik Language*. Ms. Cologne: Universität zu Köln.

——1990. The Dative in Ik and Kanuri. In W. Croft, K. Denning, and S. Kemmer (eds.). *Studies in Typology and Diachrony: Papers presented to Joseph H. Greenberg on his 75th birthday.* Typological Studies in Language 20. Amsterdam/Philadelphia: John Benjamins. 129–49.

——1997b. *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press.

HEINE, BERND. 1999. *Ik-dictionary*. Nilo-Saharan, 15. Cologne: Rüdiger Köppe.

—— 2002. Ways of Explaining Possession. In I. Baron, M. Herslund, and F. Sørensen (eds.). *Dimensions of Possession*. Amsterdam: John Benjamins. 311–28.

—— 2003 Grammaticalization. In B. D. Joseph and R. D. Janda (eds.). *The Handbook of Historical Linguistic*s. Oxford: Blackwell. 575–601.

—— and U. CLAUDI. 1986. *On the Rise of Grammatical Categories. Some Examples from Maa*. Berlin: Dietrich Reimer.

—— and T. KUTEVA. 2001. Attributive Possession in Creoles. In *Pidgins and Creoles Archive 7.* [http://www.pca.uni-siegen.de]

—— —— 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambrige University Press.

—— —— 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

—— —— 2006. *The Changing Languages of Europe.* Oxford: Oxford University Press.

—— and R. VOßEN. 1975–6. Zur Stellung der Ongamo-Sprache (Kilimandscharo). In *Afrika und Übersee* 59, 2. 81–105.

—— U. CLAUDI, and F. HÜNNEMEYER. 1991. *Grammaticalization. A Conceptual Framework*. Chicago: Chicago University Press.

HELBIG, G. and J. BUSCHA. 1996. *Deutsche Grammatik. Ein Handbuch für Ausländer*. 17th edition. Leipzig: Langenscheidt.

HEMFORTH, B. 1993. *Kognitives Parsing: Repräsentation und Verarbeitung grammatischen Wissens*. Sankt Augustin: Infix Verlag.

HENDRIKS, P. and H. DE HOOP. 2001. Optimality Theoretic Semantics. In *Linguistics and Philosophy* 24. 1–32.

HENKELMANN, P. (2006). *Die Funktionen von Akkusativmarkern aus typologischer und diachroner Sicht*. Munich: M-Press Meidenbauer.

HENTSCHEL, G. 1991. Rol' sxem formal'noj differenciacii v istoričeskom razvitii flektivnoj sistemy russkogo suščestvitel'nogo [The role of the schemes of formal differentiation in the historical development of the inflectional system of the Russian substantive]. In *Russian Linguistics* 15 (1). 31–51.

HERMON, G. 2001. Non-canonically Marked A/S in Imbabura Quechua. In A. Aikhenvald, R. M. W. Dixon, and M. Onishi (eds.). *Non-Canonical Marking of Subjects and Objects*. Amsterdam: John Benjamins. 149–77.

HETZRON, R. 1970. Toward an Amharic Case-Grammar. In *Studies in African Linguistics* 1. 301–54.

——1995. Genitival Agreement in Awngi: Variation on an Afroasiatic Theme. In F. Plank (ed.). *Double Case: Agreement by Suffixaufnahme*. Oxford: Oxford University Press. 325–35.

VON HEUSINGER, K. and J. KORNFILT. 2005. The Case of the Direct Object in Turkish: Semantics, Syntax and Morphology. In *Turkic Languages* 9. 3–44.

HEWITT, B. G. 1979. *Abkhaz*. Lingua Descriptive Studies, vol. 2. North Holland: Amsterdam.

HEWITT, G. 1996. *Georgian: A Learner's Grammar*. London: Routledge.

——2005. North West Caucasian. In *Lingua* 115. 91–145.

HEYCOCK, C. 1993. Syntactic Predication in Japanese. In *Journal of East Asian Linguistics* 2/2. 167–211.

HILDEBRANDT, K. A. 2004. A Grammar and Glossary of the Manange Language. In C. Genetti (ed.). *Tibeto-Burman Languages of Nepal: Manange and Sherpa*. Pacific Linguistics 557. Canberra: The Australian National University. 1–189.

HINRICHS, U. (ed.) 1999. *Handbuch der Südosteuropa-Linguistik*. Wiesbaden: Harrassowitz Verlag.

——2000. Prolegomena zu einer Theorie des Analytismus I, II. In U. Hinrichs and U. Büttner (eds.). *Die Südosteuropa-Wissenschaften im neuen Jahrhundert*. Wiesbaden: Harrassowitz Verlag. 83–128.

HIRAIWA, K. 2005. *Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture*. PhD Dissertation, MIT, Cambridge, MA.

HJELMSLEV, L. 1935. *La catégorie des cas: Etude de grammaire générale I*. Copenhagen: Munksgaard. [Acta Jutlandica: Aarsskrift for Aarhus Universitet 7.1.]

——1937. *La catégorie des cas: Etude de grammaire générale II*. Copenhagen: Munksgaard. [Acta Jutlandica: Aarsskrift for Aarhus Universitet 9.3.]

HOBERG, U. 1981. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache*. Munich: Hueber.

HOCKETT, C. F. 1954. Two Models of Grammatical Description. In *Word* 10. 210–31.

HODDINOTT, W. G and F. M. KOFOD. 1976. Ergative, Locative and Instrumental Case Inflections–Ngangikurungur. In R. M. W. Dixon (ed.). *Grammatical Categories in Australian Languages*. Linguistic Series No. 22. Canberra: Australian Institute of Aboriginal Studies/New Jersey: Humanities Press. 401–5.

HÖFTMANN, H. 1993. *Grammatik des Fon*. Leipzig: Langenscheidt Enzyklopädie.

HOLISKY, D. A. 1987. The Case of the Intransitive Subject in Tsova-Tush (Batsbi). In *Lingua* 71. 103–32.

HOLITSKY, D. A. 1994. (with R. Gagua) Tsova-Tush (Batsbi). In R. Smeets (ed.). *Indigenous Languages of the Caucasus: North East Caucasian Languages*. Delmar, NY: Caravan Books. 147–212.

HOLM, S. 1979. A Simple Sequentially Rejective Multiple Test Procedure. In *Scandinavian Journal of Statistics* 6. 65–70.

HOLST, J. H. 2005. *Einführung in die eskimo-aleutischen Sprachen*. Hamburg: Buske.

HOLTON, D., P. MACKRIDGE, and I. PHILIPPAKI-WARBURTON. 1997. *Greek. A Comprehensive Grammar of the Modern Language*. London: Routledge.

DE HOOP, H. 1989. Case Assignment and Generalized Quantifiers. In *Proceedings of NELS* 19. 176–90.

——1992. *Case Configuration and Noun Phrase Interpretation*. PhD Dissertation, University of Groningen.

DE HOOP, H. 1996. *Case Configuration and Noun Phrase Interpretation*. New York: Garland.
—— 1999. Optimal Case Assignment. In R. van Bezooijen and R. Kager (eds.). *Linguistics in the Netherlands*. Amsterdam/Philadelphia: John Benjamins. 97–109.
—— 2006. *On (in)animate noun phrases*. Manuscript Radboud University Nijmegen.
—— and M. J. A. LAMERS. 2006. Incremental Distinguishability of Subject and Object. In L. Kulikov, A. Malchukov, and P. de Swart (eds.). *Case, Valency, and Transitivity*. Amsterdam: John Benjamins. 269–87.
—— and A. L. MALCHUKOV. 2006. *Case Marking Strategies*. Unpublished paper. Radboud University Nijmegen.
—— —— 2007. On Fluid Differential Case Marking: a Bidirectional OT Approach. In *Lingua* 117 (Special issue on Recent Developments in Optimality Theory; Vieri Samec-Lodovici, ed.). 1636–56.
—— and B. NARASIMHAN. 2005. Differential Case-marking in Hindi. In M. Amberber and H. de Hoop (eds.). *Competition and Variation in Natural Languages: The Case for Case*. Oxford: Elsevier. 321–45.
—— —— (2008). Ergative Case-marking in Hindi. In H. de Hoop and P. de Swart (eds.). *Differential Subject Marking*. Dordrecht: Springer. 63–78.
HOPF, J.-M., J. BAYER, M. BADER, and M. MENG. 1998. Event Related Brain Potentials and Case Information in Syntactic Ambiguities. In *Journal of Cognitive Neuroscience* 10. 264–80.
—— M. BADER, M. MENG, and J. BAYER. 2003. Is Human Sentence Parsing Serial or Parallel? Evidence from Event-Related Brain Potentials. In *Cognitive Brain Research* 15. 165–77.
HOPPER, P. and S. THOMPSON. 1980. Transitivity in Grammar and Discourse. In *Language* 56. 251–99.
—— —— (eds.). 1982. *Studies in Transitivity*. Syntax and Semantics 15. New York: Academic Press.
—— and E. C. TRAUGOTT. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
HORN, L. 1984. Towards a New Taxonomy of Pragmatic Inference: Q-based and R-based Implicature. In D. Schiffrin (ed.). *Meaning, Form, and Use in Context: Linguistic Applications*. Washington: Georgetown University Press. 11–42.
HORNSTEIN, N. 2003. On Control. In R. Hendrick (ed.). *Minimalist Syntax*. Oxford: Blackwell. 6–81.
HOSOKAWA, K. 1991. *The Yawuru Language of West Kimberley: A Meaning-based Description*. PhD Thesis. Canberra: Australian National University.
HOVDHAUGEN, E., I. HOËM, C. M. IOSEFO, and A. M. VONEN. 1989. *A Handbook of the Tokelau Language*. Oslo: Norwegian University Press.
HUALDE, J. I. 2003. Case and Number Inflection of Noun Phrases. In J. I. Hualde and J. Ortiz de Urbina (eds.). *A Grammar of Basque*. Berlin: Mouton de Gruyter. 171–86.
—— and J. ORTIZ DE URBINA (eds.). 2003. *A Grammar of Basque*. Berlin/New York: Mouton.
HUANG, C.-T. J. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. Dissertation, MIT, Cambridge, MA.
HUBER, W., K. POECK, D. WENIGER, and K. WILLMES. 1983. *Der Aachener Aphasie Test (AAT)*. Göttingen: Hogrefe.
HUDSON, R. 1995. Does English Really Have Case? In *Journal of Linguistics 31*. 375–92.
HUME, A. 1612. *Prima elementa grammaticae in usum iuventutis Scoticae*. Edinburgh.

Ibragimov, G. Kh. 1978. *Rutul'skij jazyk*. Moscow: Nauka.
——1990. *Tsakhurskij jazyk*. Moscow: Nauka.
Iggesen, O. A. 2005a. Asymmetrical Case Marking, in M. Haspelmath et al. (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 206–9.
——2005b. *Case-Asymmetry. A World-Wide Typological Study on Lexeme-Class-Dependent Deviations in Morphological Case Inventories*. Munich: Lincom Europa.
——2005c. Number of Cases. In M. Haspelmath et al. (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press.
Ikegami, Y. 1981. *Suru to naru no gengogaku* [Linguistic studies on do-languages and become-languages]. Tokyo: Taishūkan.
Indefrey, P. and W. J. M. Levelt. 2004. The Spatial and Temporal Signatures of Word Production Components. In *Cognition* 92. 101–4.
Inoue, A. and J. D. Fodor. 1995. Information-Paced Parsing of Japanese. In R. Mazuka and N. Nagai (eds.). *Japanese Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum. 9–63.
Inoue, K. 2006. Case (With Special Reference to Japanese). In M. Everaert and H. van Rijmsdijk (eds.). *Blackwell Companion to Syntax*. Vol. 1. 295–374.
Insler, S. 1991–92. The Prakrit Ablative in -āhi. In *Annals of the Bhandarkar Oriental Research Institute* 72–73. 15–21.
Iordanidi, S. I. and V. B. Krys'ko. 2000. *Istoričeskaja grammatika drevnerusskogo jazyka*. Tom 1. *Množestvennoe čislo imennogo sklonenija*. [A historical grammar of Old Russian. Vol. 1. Nominal declension: Plural]. Moscow: Azbukovnik.
Ivanov, V. V. 1983. *Istoričeskaja grammatika russkogo jazyka.* [A historical grammar of Russian]. Moscow: Prosveščenie.
Ivić, P. 2001. Balkansprachbund und Sprachgeographie. In H. Schaller (ed.). *Grundfragen eines Südosteuropa-Sprachatlas. Geschichte. Problematik. Method*. Pilotprojekt. Marburg: Biblion Verlag. 45–52.
Iwasaki, S. 1987. Identifiability, Scope-Setting, and the Particle 'wa': A Study of Japanese Spoken Expository Discourse. In J. Hinds et al. (eds.). *Perspectives on Topicalization–The Case of Japanese 'wa'*. Amsterdam: John Benjamins. 107–41.
——2002. *Japanese*. Amsterdam: John Benjamins.
Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: The MIT Press.
——1976. Toward an Explanatory Semantic Representation. In *Linguistic Inquiry* 7(1). 89–150.
——1983. *Semantics and Cognition*. Cambridge, MA: MIT Press.
——1987. The Status of Thematic Relations in Linguistic Theory. In *Linguistic Inquiry* 18(3). 369–411.
——1990. *Semantic Structures*. Cambridge, MA: The MIT Press.
——and B. Landau. 1992. Spatial Language and Spatial Cognition. In R. Jackendoff (ed.). *Languages of the Mind: Essays on Mental Representation*. Cambridge, MA: MIT. 99–124.
Jacobsen, W. 1964. *A Grammar of the Washo Language*. PhD Diss. University of California at Berkeley.
Jacobsen, W. M. 1992. *The Transitive Structure of Events in Japanese.* Tokyo: Kurosio Publishers.
Jaeger, F. T. and R. Levy. 2006. *The Structural Basis of Probabilistic Syntactic Production*. Paper presented at the 12th Annual Conference on Architectures and Mechanisms for Language Processing, Nijmegen, The Netherlands.

Jäger, G. 2004. Learning Constraint Sub-hierarchies. The Bidirectional Gradual Learning Algorithm. In R. Blutner and H. Zeevat (eds.). *Optimality Theory and Pragmatics*. Houndmills: Palgrave MacMillan. 251–87.

—— and H. Zeevat. 2002. A Statistical Reinterpretation of Harmonic Alignment. In *Proceedings of the 4th Tbilisi Symposium on Logic, Language and Linguistics*. 173–81.

Jahani, C. 2003. The Case System in Iranian Balochi in a Contact Linguistic Perspective. In C. Jahani and A. Korn (eds.). *The Baloch and Their Neighbours*. Wiesbaden: Reichert.

Jake, J. 1985. *Grammatical Relations in Imbambura Quechua*. New York: Garland.

Jakobson, R. 1931. K xarakteristike èvrazijskogo jazykovogo sojuza. In *Selected Writings* 1. The Hague: Mouton [1970]. 144–201.

—— 1936. Beiträge zur Allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus. In *Travaux du Cercle Linguistique de Prague* 6. 240–88. [Reprinted in 1966 in E. P. Hamp, F. W. Householder, and R. Austerlitz (eds.). *Readings in Linguistics* II. Chicago: The University of Chicago Press. 51–89. Reprented in 1971 in *Selected Writings II*. Den Haag: Mouton. 23–72.]

Jakobson, R. 1958 [1971]. Morphological inquiry into Slavic declension: structure of Russian case forms [text in Russian, summary in English in 1971 version]. In R. Jakobson. *Selected writings*, vol. II: Words and Language. The Hague: Mouton. 154–83.

Janda, L. 1993. A Geography of Case Semantics: *The Czech Dative and the Russian Instrumental*. Berlin/New York: Mouton de Gruyter.

Janssen, D., B. Bickel, and F. Zúñiga. 2006. Randomization Tests in Language Typology. In *Linguistic Typology* 10. 419–40.

Jelaska, Z., M. Kovačević, and M. Anđel. 2002. 'Morphology and Semantics – The Basis of Croatian Case', in M. D, Voeikova, and W. U. Dressler (eds.). *Pre- and Protomorphology: Early Phases of Morphological Development in Nouns and Verbs*. München: LINCOM Europa, 177–89.

Jelinek, E. 1984. Enpty Categories, Case and Configurationality. In *Natural Language and Linguistic Theory* 2. 39–76.

Jensen, J. T. 1977. *Yapese Reference Grammar*. Honolulu: University of Hawai'i Press.

Jespersen, O. 1924. *The Philosophy of Grammar*. London: G. Allen & Unwin.

Johanson, L. 2002. *Structural Factors in Turkic Language Contacts*. London: Curzon.

—— 2006. Two Approaches to Specificity. In L. Kulikov, A. Malchukov, and P. de Swart (eds.). *Case, Valency and Transitivity*. Amsterdam and Philadelphia: John Benjamins. 123–35.

Johns, A., D. Massam, and J. Ndayiragije. 2006. *Ergativity: Emerging Issues*. Dordrecht: Springer.

Johnson, K. 1991. Object Positions. In *Natural Language and Linguistic Theory* 9. 577–636.

Jones, M. A. 1988. Cognate Objects and the Case Filter. In *Journal of Linguistics* 24. 89–110.

Jones, M. C. 2002. Mette a haout dauve la grippe des Angllaïs: Convergence on the Island of Guernsey. In M. C. Jones, and E. Esch (eds.). *Language Change. The Interplay of Internal, External and Extra-linguistic Factors*. Berlin/New York: Mouton de Gruyter. 143–68.

Joshi, A. K. 1985. How Much Context-sensitivity is Required to Provide Reasonable Structural Descriptions: Tree Adjoining Grammars. In D. Dowty, L. Kartunnen, and A. Zwicky (eds.). *Natural Language Parsing: Psycholinguistic, Computational and Theoretical Perspectives*. Cambridge: Cambridge University Press.

—— L. Levy, and M. Takahashi. 1975. Tree Adjunct Grammars. In *Journal of the Computer and System Sciences* 10. 136–63.

JUNKER, M.-O. 2003. A Native American View of the 'mind' as seen in the Lexicon of Cognition in East Cree. In *Cognitive Linguistics* 14(2–3). 167–94.
—— In press. Semantic Primes and their Grammar in a Polysynthetic Language: East Cree. In C. Goddard (ed.). *Studies in Cross-Linguistic Semantics*. Amsterdam: John Benjamins.
—— and LOUISE BLACKSMITH. 2006. Are There Emotional Universals? Evidence from the Native American language East Cree. In *Culture & Psychology* (12)3. 275–304.
JURAFSKY, D. 2003. Probabilistic Modeling in Psycholinguistics: Linguistic Comprehension and Production. In R. Bod, J. Hay, and S. Jannedy (eds.). *Probabilistic Linguistics.* Cambridge, MA: MIT Press. 39–95.
KAAN, E., F. WIJNEN, and T. Y. SWAAB. 2004. Gapping: Electrophysiological Evidence for Immediate Processing of 'Missing' Verbs in Sentence Comprehension. In *Brain and Language* 89. 584–92.
KACHRU, Y. 1987. Ergativity, Subjecthood and Topicality in Hindi-Urdu. In *Lingua* 71. 223–38.
KAHR, J. C. 1975. Adpositions and Locationals: Typology and Diachronic Development. *Working Papers on Language Universals (Stanford)* 19: 21–54.
—— 1976. The Renewal of Case Morphology: Sources and Constraints. In *Working Papers on Language Universals* (Stanford University) 20. 107–51.
KÄMPFE, H.-R. and A. P. VOLODIN. 1995. *Abriß der tschuktschischen Grammatik auf der Basis der Schriftsprache*. Tunguso Sibirica, 1. Wiesbaden: Harrassowitz.
KARLSSON, F. 1982. *Suomen Peruskielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura.
—— 1983. *Finnish: An Essential Grammar*. Juva: WSOY.
—— 1999. *Finnish. An Essential Grammar*. London: Routledge.
KASTENHOLZ, R. 1989. *Grundkurs Bambara (Manding).* Cologne: Köppe.
KEENAN, E. L. 1974. The Functional Principle: Generalizing the Notion of 'Subject of'. In *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 298–309.
KEENAN, E. L. 1989. Semantic Case Theory. In R. Bartsch, J. van Benthem, and P. van Emde Boas (eds.). *Semantics and Contextual Expression*. Dordrecht: Foris. 33–56.
—— and B. COMRIE. 1977. Noun Phrase Accessibility Hierarchy and Universal Grammar. In *Linguistic Inquiry* 8. 63–99.
VAN KEMENADE, A. 1987. *Syntactic Case and Morphological Case in the History of English*. PhD dissertation, Utrecht University.
KEMMER, S. 1993. *The Middle Voice*. Amsterdam: John Benjamins.
KEMPE, V. and B. MACWHINNEY. 1998. The Acquisition of Case Marking by Adult Learners of Russian and German. In *Studies in Second Language Acquisition* 20, N 4. 543–87.
KEMPEN, G. 1977. Conceptualizing and Formulating in Sentence Production. In S. Rosenberg (ed.). *Sentence Production: Developments in Research and Theory*. Hillsdale, NJ: Erlbaum. 259–75.
—— and K. HARBUSCH. 2002. An Artificial Opposition between Grammaticality and Frequency: Comment on Bornkessel, Schlesewsky and Friederici (2002). In *Cognition* 90. 205–10.
—— and E. HOENKAMP. 1987. An Incremental Procedural Grammar for Sentence Formulation. In *Cognitive Science* 11. 201–58.
—— and P. HUIJBERS. 1983. The Lexicalization Process in Sentence Production and Naming: Indirect Election of Words. In *Cognition* 14. 185–209.
KENESEI, I., R. M. VAGO, and A. FENYVESI. 1998. *Hungarian*. London: Routledge.

978–0–19–920647–6 59-Spencer-Ref OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 856 of 894 May 16, 2008 19:1

Kerimova, A. A. and V. S. Rastorgueva. 1975. Kategorija padezha [The category of case]. In V. S. Rastorgueva (ed.). *Opyt istoriko-tipologicheskogo issledovanija iranskix jazykov* [A Diachronic-Typological Study of the Iranian Languages.] Vol. 2. Moscow: Nauka. 117–99.

van de Kerke, S. 1996. *Affix Order and Interpretation in Bolivian Quechua*. PhD dissertation, University of Amsterdam.

Kibrik, A. E. 1970. *K tipologii prostranstvennykh znachenij (na materiale dagestanskikh jazykov)*. In *Jazyk i chelovek*. Moscow: MGU.

——1979. Canonical Ergativity and Daghestan Languages. In F. Plank (ed.). *Ergativity*. London: Academic Press. 61–77.

——1985. Toward a Typology of Ergativity. In J. Nichols and A. Woodbury (eds.). *Grammar Inside and Outside the Clause*. Cambridge: Cambridge University Press. 268–324.

——1991. Organizing Principles for Nominal Paradigms in Daghestanian Languages. In F. Plank (ed.). *Paradigms: The Economy of Inflection*. Berlin: Mouton de Gruyter. 225–74.

——1992. Principy organizacii imennoj paradigmy v dagestanskikh jazykakh (sopostavitel'no-tipologicheskie nabl'udenija). In A. E. Kibrik. *Ocherki po obshchim i prikladnym voprosam jazykoznanija*. Moscow: MGU.

—— and D. S. Sanedor. 1994a. Archi. In R. Smeets (eds.). *The Indigenous Languages of the Caucasus*, vol. 4. Delmar, New York: Caravan Books. 297–365.

———1994b. Khinalug. In R. Smeets (ed.). *The Indigenous Languages of the Caucasus*, vol. 4. Delmar, New York: Caravan Books. 368–407.

——1995. Direct–oblique agreement of attributes in Tsez and elsewhere in Daghestanian. In F. Plank (ed.). *Double Case: Agreement by Suffixaufnahme*. New York/Oxford: Oxford University Press. 216–39.

——S. G. Tatevosov, and A. Eulenberg. (eds.). 1996. *Godoberi.* Munich/Newcastle: Lincom Europa.

——1997. Beyond Subject and Object: Towards a Comprehensive Relational Typology. In *Linguistic Typology* 1(3). 279–346.

——1998. Archi (Caucasian–Daghestanian). In A. Spencer and A. Zwicky (eds.). *Handbook of Morphology*. Oxford: Blackwell Publishers. 455–76.

—— and Ja. G. Testelets. (eds.). 1999. *Elementy caxurskogo jazyka v tipologičeskom osveščenii*. Moscow: Nasledie.

——K. I. Kazenin, E. A. Ljutikova, and S. G. Tatevosov. (ed.). 2001. *Bagvalinskij jazyk. Grammatika. Teksty. Slovari*. [The Bagval language. Grammar. Texts. Lexica.]. Moscow: Nasledie.

——2003a. Nominal Inflection Galore: Daghestanian, with side glances at Europe and the World. In: F. Plank (ed.). *Noun Phrase Structure in the Languages of Europe*. Mouton de Gruyter.

——2003b. *Konstanty i peremennyje jazyka.* Moscow: Aleteia.

——What's in the Head of Head-Marking Languages or Is there Case in Head-Marking Languages? In B. Comrie and V. Solovyev (eds.). *The Typology of Argument Structure and Grammatical Relations*. Amsterdam: John Benjamins.

—— and S. Kodzasov. 1990. *Sopostavitel'noe izuchenie dagestanskix jazykov. Im'a. Fonetika*. Moscow: MGU.

—— and Ja. G. Testelec (eds.). 1999. *Èlementy caxurskogo jazyka v tipologičeskom osveščenii*. Moscow: Nasledie.

———2004. Bezhta. In M. Job (ed.). *Indigenous Languages of the Caucasus*, Vol. 3. *The North East Caucasian Languages*. Ann Arbor.

——S. V. Kodzasov, and I. P. Olovjannikova, 1972. *Fragmenty grammatiki khinalugskogo jazyka*. Moscow: MGU.

———————1977. *Opyt strukturnogo opisanija arčinskogo jazyka* [Towards a structural description of Archi]. (In Russian). Vols. 1–4. Moscow: Izdatel'stvo Moskovskogo Universiteta.

———and I. A Muravjova. 2000. *Jazyk i fol'klor aljutorcev*. Moscow: Nasledie. [English translation, 2004. Language and Folklore of the Alutor People. Osaka: Endangered Languages of the Pacific Rim.]

——K. I. Kazenin, E. A. Ljutikova, and S. G. Tatevosov (eds.). 2001. *Bagvalinskij jazyk: gramatika, teksty, slovari*. Moscow: Nasledie.

——S. V. Kodzasov and I. A. Muravyova. 2004. *Language and Folklore of the Alutor People*. Suita, Japan: ELPR Publications.

Kiefer, F. 1987. The Cases of Hungarian Nouns. In *Acta Linguistica Academiae Scientiarum Hungaricae*, vol. 37. 93–101.

Kießling, R. 1994. *Eine Grammatik des Burunge*. Afrikanistische Forschungen 13. Hamburg: Research and Progress.

——2001. *Case in Datooga*. Ms.

Kilani–Schoch, M., and Dressler, W. U. 2002. The Emergence of Inflectional Paradigms in Two French Corpora: An Illustration of General Problems of Pre- and Protomorphology. In: M. D. Voeikova and W. U. Dressler (eds.). *Pre- and Protomorphology: Early Phases of Morphological Development in Nouns and Verbs*. Lincom Studies in Theoretical Linguistics 29. Munich: Lincom Europa. 45–61.

——A. de Marco, A. Christofidou, M. Vassilakou, R. Vollmann, and W. U. Dressler. 1997. On the Demarcation of Phases in Early Morphology Acquisition in Four Languages. In K. Dziubalska-Kołaczyk (ed.). *Pre- and Protomorphology in Language Acquisition*. Papers and Studies in Contrastive Linguistics V. 33. Poznań: Adam Mickiewicz University. 15–33.

Kilby, D. 1981. On Case Markers. In *Lingua* 54. 101–33.

Kim, S. and J. Maling. 1993. Syntactic Case and Frequency Adverbials in Korean. In *Harvard Studies in Korean Linguistics,* V. 368–78.

———————1998. Case Assignment in the *siphta*-construction. In R. King (ed.). *Description and Explanation in Korean Linguistics*. Cornell East Asian Series. Ithaca, NY: The Cornell University Press. 133–68.

Kim, Y. 1999. The Effects of Case Marking Information on Korean Sentence Processing. In *Language and Cognitive Processes* 14(5/6). 687–714.

Kimball, G. D. 1985. *A Descriptive Grammar of Koasati*. Dissertation, Tulane University, New Orleans.

——1991. *Koasati Grammar: Studies in the Anthropology of North American Indians*. Lincoln/London: University of Nebraska Press.

Kimball, J. 1973. Seven Principles of Surface Structure Parsing in Natural Language. In *Cognition* 2. 15–47.

Kinsui, S. 1995. Katari no 'wa' ni kansuru oboegaki [Memoranda on ~*wa* in Old Japanese speech]. In T. Masuoka et al. (eds.). *Nihongo no shudai to toritate* [Topic and topicalization in Japanese]. Tokyo: Kuroshio. 71–80.

Kiparsky, P. 1973. 'Elsewhere' in Phonology. In S. R. Anderson and P. Kiparsky (eds.). *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston. 93–106.

——1987. *Morphology and Grammatical Relations*. Unpublished ms., Stanford University.

KIPARSKY, P. 1988. *Agreement and Linking Theory*. Stanford University.
——1997. The Rise of Positional Licensing. In A. van Kemenade and N. Vincent (eds.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press. 460–94.
——1998. Partitive Case and Aspect. In M. Butt and W. Geuder (eds.). *The Projection of Arguments: Lexical and Compositional Factors*. Stanford: CSLI Publications. 259–307.
——2001. Structural Case in Finnish. In *Lingua* 111. 315–76.
KIPARSKY, V. 1967. *Russische historische Grammatik*, Bd. 2: *Die Entwicklung des Formensystems*. Heidelberg: Winter.
KISS, K. 2000. Effect of Verb Complexity on Agrammatic Aphasics' Sentence Production. In R. Bastiaanse and Y. Grodzinsky (eds.). *Grammatical Disorders in Aphasia*. London: Whurr Publishers. 123–51.
KITTILÄ, S. 2002. *Transitivity: Towards a Comprehensive Typology*. Publications in General Linguistics 5. Turku: Åbo Akademis Tryckeri.
——2005a. *Differential Goal Marking: On the Encoding of Animate and Inanimate Goals*. Unpublished paper, University of Turku.
——2005b. A Typology of Involuntary Agent Constructions. In *Word* 56.3.
——2006a. Object-, Animacy- and Role-based Strategies. A Typology of Object Marking. In *Studies in Language* 30. 1–32.
——2006b. The Woman Showed the Baby to her Sister. On Resolving Animacy-driven Ambiguity in Ditransitives. In L. Kulikov, A. Malchukov, and P. De Swart (eds.). *Studies in Case, Valency and Transitivity*. Amsterdam: Benjamins. 291–309.
KLAIMAN, M. H. 1980. Bengali Dative Subjects. In *Lingua* 51. 275–95.
——1987. Mechanisms of Ergativity in South Asia. In *Lingua* 71. 61–102.
KLAMPFER, S. and K. KORECKY-KRÖLL. 2002. Nouns and Verbs at the Transition from Pre- to Protomorphology: A Longitudinal Case Study on Austrian German. In M. D. Voeikova, and W. U. Dressler (eds.). *Pre- and Protomorphology: Early Phases of Morphological Development in Nouns and Verbs*. LINCOM Studies in Theoretical Linguistics 29. Munich: LINCOM Europa. 61–75.
KLEIST, K. 1905. Über Leitungsaphasie. In *Monatschrift für Neurologie und Psychiatrie* 17. 503–32.
KLENIN, E. 1980. Conditions on Object Marking: Stages in the History of the East Slavic Genitive-accusative. In E. Traugott, R. Labrum, and S. Shepherd (eds.). *Papers from the 4th International Conference on Historical Linguistics*. Amsterdam: Benjamins. 253–7.
KLIMOV, G. A. 1962. *Sklonenie v kartvel'skix jazykax v sravnitel'no-istoričeskom aspekte* [Declension in Kartvelian languages in a comparative-historical perspective]. Moscow: Izdatel'stvo Akademii Nauk SSSR.
VAN KLINKEN, C. L. 1999. *A Grammar of the Fehan Dialect of Tetun. An Austronesian Language*. Canberra: Pacific Linguistics.
KNOEFERLE, P. and M. W. CROCKER. 2006. The Coordinated Interplay of Scene, Utterance, and World Knowledge: Evidence from Eye Tracking. In *Cognitive Science* 30. 481–529.
KOBAYASHI, J. 1984. Nihongo no joshi 'wa', 'ga' to birumago no joshi 'ha', 'ka' no taishō kenkyū [A contrastive study on the particles ∼*wa* and ∼*ga* in Japanese and ∼*ha* and ∼*ka* in Burmese]. In: *Nihongo kyōiku* [Education of the Japanese language] 54. 89–98.
KOEHN, E. and S. KOEHN. 1986. Apalai. In D. Derbyshire and G. K. Pullum (eds.). *Handbook of Amazonian Languages*, vol. 1. Berlin: Mouton de Gruyter. 33–127.
KOLK, H. 1995. A Time-based Approach to Agrammatic Production. In *Brain and Language* 50. 282–303.

978–0–19–920647–6 59-Spencer-Ref OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 859 of 894 May 16, 2008 19:1

——and A. D. Friederici. 1985. Strategy and Impairment in Sentence Understanding by Broca's and Wernicke's Aphasics. In *Cortex* 21. 47–67.

——and M. Van Grunsven. 1985. Agrammatism as a Variable Phenomenon. In *Cognitive Neuropsychology* 2. 347–84.

——D. J. Chwilla, M. Van Herten, and P. J. W. Oor. 2003. Structure and Limited Capacity in Verbal Working Memory: A Study with Event-Related Potentials. In *Brain and Language* 85. 1–36.

Koneski B., B. Vidoeski, and O. Jašar-Nasteva. 1968. Distribution des balkanismes en macedonien. In *Actes du Premier Congrès international des études balkaniques et sud-est européennes*. Vol. 6. *Linguistique*. Sofia: Académie bulgar des sciences. 517–48.

Konieczny, L. 1996. *Human Sentence Processing: A Semantics-Oriented Parsing Approach* (IIG-Bericht Nr. 3/96). Unpublished Doctoral dissertation, Universität Freiburg, Freiburg i.Br.

König, C. 2002. *Kasus im Ik*. Nilo-Saharanische Reihe 16. Cologne: Rüdiger Köppe.

——2006. Marked Nominative in Africa. In *Studies in Language* 30(4). 655–732.

——(2008a.) *Case in Africa*. Oxford University Press.

——(2008b). The Marked-Nominative Languages of Eastern Africa. In B. Heine and D. Nurse. *A Linguistic Geography of Africa*. Cambridge: Cambridge University Press. 251–71.

Konopka, A. and K. Bock. 2005. *Helping Syntax out: How much do Words do?* Paper presented at the 18th Annual CUNY Conference on Human Language Processing, Tucson, Arizona. *Konstanz Universals Archive*. http://ling.uni-konstanz.de/pages/proj/sprachbau.htm

Koopman, H. 2006. Agreement Configurations: In Defense of 'Spec head'. In C. Boeckx (ed.). *Agreement Systems*. Amsterdam: Benjamins. 159–99.

——and D. Sportiche. 1991. The Position of Subjects. In *Lingua* 85. 211–58.

Koopman, W. 1990. *Word Order in Old English*. PhD dissertation, University of Amsterdam.

——1993. The Order of Dative and Accusative Objects in Old English and Scrambling. In *Studia Anglica Posnaniensia* xxv–xxvii. 109–21.

Koptjevskaja-Tamm, M. 1993. *Nominalizations*. London: Routledge.

——2000. Romani Genitives in Cross-linguistic Perspective. In V. Elšík and Y. Matras (eds.). *Grammatical Relations in Romani: The Noun Phrase*. Amsterdam: John Benjamins. 123–49.

——2003a. Possessive Noun Phrases in the Languages of Europe. In F. Plank (ed.). *Noun Phrase Structure in the Languages of Europe*. Berlin: Mouton de Gruyter. 621–722.

——2003b. Action Nominal Constructions in the Languages of Europe. In F. Plank (ed.). *Noun Phrase Structure in the Languages of Europe*. Berlin: Mouton de Gruyter. 723–59.

——2004. *Maria's Ring of gold*: Adnominal Possession and Non-anchoring Relations in European Languages. In J. Kim, Y. A. Lander, and B. H. Partee (eds.). *Possessives and Beyond: Semantics and Syntax*. Amherst, MA: GLSA Publications. 155–81.

Korn, A. (forthcoming). How Patient is a Definite Patient? Marking of Arguments in Balochi Ergative and Mixed Constructions. In S. Karimi et al. (eds.). *Proceedings of the First International Conference in Aspects of Iranian Linguistics*. Newcastle-upon-Tyne: Cambridge Scholars Press.

Kornfilt, J. 2006. *DOM and Two Types of DSM in Turkish*. Ms., Syracuse University.

Kortlandt, F. H. H. 1983. On Final Syllables in Slavic. In *Journal of Indo-European Studies* 11. 167–85.

KORTLANDT, F. H. H. 1984. Proto-Armenian Case Endings. In *International Symposium on Armenian Linguistics: Reports* [Meždunarodnyj simpozium po armjanskomu jazykoznaniju: Doklady], 97–106. Erevan: AN Arm. SSR. [Reprinted in F. H. H. Kortlandt. 2003. 45–51.]

—— 2003. *Armeniaca: Comparative Notes*. Ann Arbor: Caravan.

—— 2005. On the Accentuation of the Illative. In *Baltu Filologija* 14 (1). 67–9.

KOSHAL, S. 1979. *Ladakhi Grammar*. Delhi: Motilal Banarsidass.

KOTZ, S. A., S. FRISCH, D. Y. VON CRAMON, and A. D. FRIEDERICI. 2003. Syntactic Language Processing: ERP Lesion Data on the Role of the Basal Ganglia. In *Journal of the International Neuropsychological Society* 9. 1053–60.

KOZINSKY, I., V. P. NEDJALKOV, and M. S. POLINSKAJA. 1988. Antipassive in Chukchee: Oblique Object, Object Incorporation, Zero Object. In M. Shibatani (ed.). *Passive and Voice*. Amsterdam: John Benjamins. 651–706.

KOZINTSEVA, N. 1995. *Modern Eastern Armenian*. Munich/Newcastle: Lincom Europa.

KRACHT, M. 2002. On the Semantics of Locatives. In *Linguistics and Philosophy* 25. 157–232.

—— 2003. Against the Feature Bundle Theory of Case. In E. Brandner, and H. Zinsmeister (eds.). *New Perspectives on Case Theory*. Stanford: CSLI. 165–90.

—— 2005. The Semantics of Locatives in the Uralic Languages. In J. Fernandez-Vest (ed.). *Les langues ouraliennes aujourd'hui: approche linguistique et cognitive*. Paris: Honoré Champion. 145–58.

KREJNOVIČ, E. A. 1958. *Jukagirskij jazyk* [The Yukaghir language]. Moscow: Nauka.

—— 1982. *Issledovanija i materialy po jukagirskomu jazyku*. [The Yukaghir language: studies and materials.] Moscow/Leningrad: Nauka.

KRIFKA, M. 1989. *Nominalreferenz und Zeitkonstitution: Zur Semantik von Massentermen, Pluraltermen und Aktionsarten*. Munich: Fink.

—— 1992. Thematic Relations as Links between Nominal Reference and Temporal Constitution. In I. Sag and A. Szabolcsi (eds.). *Lexical Matters*. Stanford: CSLI. 29–53.

—— 2004. Semantic and Pragmatic Conditions for the Dative Alternation. In *Korean Journal of English Language and Linguistics* 4. 1–32.

KRISHNAMURTI, B. 2003. *The Dravidian Languages*. Cambridge: Cambridge University Press.

KROCH, A. and A. K. JOSHI. 1985. *The Linguistic Relevance of Tree Adjoining Grammar*. Technical Report MS-CS-85–16, Department of Computer and Information Sciences, University of Pennsylvania.

KROEGER, P. R. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford, CA: CSLI Publications, Stanford University.

KRUSPE, N. 1999. *A Grammar of Semelai*. PhD. University of Melbourne.

KUIPERS, A. H. 1967. *The Squamish Language. Grammar, Texts, Dictionary*. The Hague/Paris: Mouton.

KULIKOV, L. 2006. Case Systems in a Diachronic Perspective: A Typological Sketch. In L. Kulikov, A. Malchukov, and P. de Swart (eds.). *Case, Valency and Transitivity*. Amsterdam: Benjamins. 23–47.

—— (to appear). Case Variation in Indo-European and Beyond: A Diachronic Typological Overview.

KUMAXOV, M. 1971. *Slovoizmenenie adygskix jazykov* [Inflectional system of the Adyghe languages]. Moscow: Nauka.

—— 1989. *Sravnitel'no-istoričeskaja grammatika adygskix (čerkesskix) Jazykov* [A comparative-historical grammar of the Adyghe (Circassian) languages]. Moscow: Nauka.

——K. Vamling, and Z. Kumakhova. 1996. Ergative Case in Circassian Languages. In *Lund University/Department of Linguistics Working Papers* 45. 93–111.

Kuno, S. 1973. *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.

Küntay, A. and D. I. Slobin. 1996. Listening to a Turkish Mother: Some Puzzles for Acquisition. In D. I. Slobin, J. Gerhardt, A. Kyratzis, and J. Guo (eds.). *Social Interaction, Social Context, and Language: Essays in Honor of Susan Ervin-Tripp*. Hillsdale, NJ: Lawrence Erlbaum Associates. 265–86.

—— ——1999. The Acquisition of Turkish as a Native Language. A Research Review. In *Turkic Languages* 3. 151–88.

—— ——2002. Putting Interaction back into Child Language: Examples from Turkish. In *Psychology of Language and Communication* 6. 5–14.

Kuroda, Sh. 2005. Focusing on the Matter of Topic: A Study of 'wa' and 'ga' in Japanese. In *Journal of East Asian Linguisitics* 14-1. 1–58.

Kuryłowicz, J. 1949. Le problème du classement des cas. In *Biuletyn Polskiego Towarzystwa Językoznawczego* 9. 20–43. [Reprinted in *Esquisses linguistiques*. Wrocław & Kraków. 131–50.]

Kutsch Lojenga, C. 1994. *Ngiti: A Central Sudanic Language of Zaire*. Cologne: Rüdiger Köppe.

Laalo, K. 2002. Acquisition of Case in Finnish: A Preliminary Overview. In M. D. Voeikova and W. U. Dressler (eds.). *Pre- and Protomorphology: Early Phases of Morphological Development in Nouns and Verbs.* Lincom Studies in Theoretical Linguistics, 29. Munich: Lincom Europa. 83–105.

Lafitte, P. 1962 [1978]. *Grammaire Basque* (Navarro-Labourdin Littéraire). Revised edition. Donostia: Elkar.

Laka, I. 1993. The Structure of Inflection: A Case Study in X°. In J. I. Hualde and J. Ortiz de Urbina (eds.). *Generative Studies in Basque Linguistics.* Amsterdam: John Benjamins. 21–70.

Lakoff, G. 1977. Linguistic Gestalts. *Papers from the Thirteenth Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 236–87.

——1987. *Women, Fire and Dangerous Things*. Chicago: Chicago University Press.

——and M. Johnson. 1980. *Metaphors We Live By*. Chicago: Chicago University Press.

Lambrecht, K. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

Lamers, M. J. A. 2001. *Sentence Processing: Using Syntactic, Semantic, and Thematic Information*. Grodil 33. University of Groningen, Groningen.

Lamers, M., and H. de Hoop. 2005. Animacy Information in Human Sentence Processing: An Incremental Optimization of Interpretation Approach. In H. Christiansen, P. R. Skadhauge, and J. Villadsen (eds.). *Constraint Solving and Language Processing (CLSP 2004)*. Heidelberg: Springer. 158–71.

Lamiroy, B. and N. Delbecque. 1998. The Possessive Dative in Romance and Germanic Languages. In W. Van Belle, and W. Van Langendonck (eds.). *The Dative*, vol. 2. Amsterdam: John Benjamins. 29–74.

Lancelot, C. and Arnauld, A. 1660. *Grammaire générale et raisonnée*. Paris.

Landau, I. 2000. *Elements of Control: Structure and Meaning in Infinitival Constructions*. Dordrecht: Kluwer.

——2002. *The Locative Syntax of Experiencers*. Ms.

Lander, Y. A. 2004. Dealing with Relevant Possessors. In J. Kim, Y. A. Lander, and B. H. Partee (eds.). *Possessives and Beyond: Semantics and Syntax*. Amherst, MA: GLSA Publications. 309–36.

Langacker, R. W. 1987. *Foundations of Cognitive Grammar*, Vol. 1: *Theoretical Prerequisites*. Stanford: Stanford University Press.

——1991. *Foundations of Cognitive Grammar*, Vol. 2: *Descriptive Application*. Stanford: Stanford University Press.

——1993. Reference-point Constructions. In *Cognitive Linguistics* 4. 1–38.

——1995. Possession and Possessive Constructions. In J. R. Taylor and R. E. MacLaury (eds.). *Language and the Cognitive Construal of the World.* Berlin: Mouton de Gruyter. 51–79.

Langdon, M. 1970. *A Grammar of Diegueño. The Mesa Grande dialect*. Berkeley: University of California Press.

van Langendonck, W. 1998. The Dative in Latin and the Indirect Object in Dutch. In W. van Langendonck and W. van Belle (eds.). *The Dative*. Volume 2: *Theoretical and Contrastive Studies*. Amsterdam/Philadelphia: John Benjamins. 211–59.

LaPolla, R. J. 1993. Arguments Against 'Subject' and 'Direct Object' as Viable Concepts in Chinese. In *Bulletin of the Institute of History and Philology* 63.4. 759–813.

Larson, R. K. 1988. On the Double Object Construction. In *Linguistic Inquiry* 19. 335–91.

Lasnik, H. and M. Saito. 1984. On the Nature of Proper Government. In *Linguistic Inquiry* 15. 235–89.

—— ——1992. *Move-α*. Cambridge, MA: MIT Press.

Last, M. and D. Lucassen. 1998. A Grammatical Sketch of Chai, a Southeastern Surmic Language. In G. J. Dimmendaal and M. Last (eds.). *Surmic Languages and Cultures.* Nilo-Saharan 13. Cologne: Rüdiger Köppe. 375–436.

Laubscher, G. G. 1921. *The Syntactical Causes of Case Reduction in Old French*. Elliot Monographs 7. Princeton, NJ: Princeton University Press.

Laughlin, C. 1975. Lexicostatistics and the Mystery of So Ethnolinguistic Relations. In *Anthropological Linguistics* 17. 325–41.

Laughren, M. 1992. Secondary Predication as a Diagnostic of Underlying Structure in Pama-Nyungan Languages. In I. Roca (ed.). *Thematic Structure: Its Role in Grammar*. Bedim Foris/Walter de Gruyter. 199–246.

Lazard, G. 1992. *A Grammar of Contemporary Persian*. Cosa Mesa: Mazda Publishers.

——1998. *Actancy*. Berlin: Mouton.

——2003. Transitivity Revisited as an Example of a More Strict Approach to Typological research. In *Folia Linguistica* XXXVI 3–4. 141–90.

Lee, H. 2001a. Markedness and Word Order Freezing. In P. Sells (ed.). *Formal and Empirical Issues in Optimality Theoretic Syntax*. Stanford: CSLI Publications.

——2001b. *Optimization in Argument Expression and Interpretation: A Unified Approach*. Doctoral dissertation, Stanford University.

——2003. Parallel Optimization in Case Systems. In M. Butt and T. H. King (eds.). *Nominals: Inside and Out*. Stanford, CA: CSLI Publications. 15–58.

——(2008). Quantitative Variation in Korean Case Ellipsis: Implications for Case Theory. In H. de Hoop and P. de Swart (eds.). *Differential Subject Marking*. Dordrecht: Springer. 41–61.

Lee, S.-H. 1990. *Konfrontative Analyse zwischen dem ungarischen und koreanischen Kasussystem*. Wiesbaden: Harrassowitz.

978–0–19–920647–6 59-Spencer-Ref OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 863 of 894 May 16, 2008 19:1

Lefebvre, C. and A. Brousseau. 2002. *A Grammar of Fongbe.* Berlin: Mouton de Gruyter.

Legate, J. A. 2005. *Morphological and Abstract Case*. Ms. University of Delaware.

Legendre, G., W. Raymond, and P. Smolensky. 1993. An Optimality-Theoretic Typology of Case and Grammatical Voice Systems. In *Proceedings of the 19th Meeting of the Berkeley Linguistics Society*. Berkeley. 464–78.

Lehiste, I. 1997. Search for Phonetic Correlates in Estonian Prosody. In I. Lehiste and J. Ross (eds.). *Estonian Prosody: Papers from a Symposium*. Tallinn: Institute of Estonian Language.

Lehmann, C. 1982. *Thoughts on Grammaticalization. A Programmatic* Sketch, Volume 1. Arbeiten des Kölner Universalien-Projekts, 48. Cologne: University of Cologne, Institut für Sprachwissenschaft.

——1985. Latin Case Relations in Typological Perspective. In C. Touratier (ed.). *Syntaxe et latin*. Aix-en-Provence: Université de Provence. 81–104.

——1988. On the Function of Agreement. In M. Barlow and A. Ferguson (eds.). *Agreement in Natural Language*. Stanford, CA: Centre for the Study of Language and Information. 55–65.

——1991. Predicate Classes and PARTICIPATION. In H. Seiler and W. Premper (eds.). *Partizipation : das sprachliche Erfassen von Sachverhalten*. Tübingen: Narr. 183–239.

——1995. *Thoughts on Grammaticalization*. Lincom Studies in Theoretical Linguistics 01. Munich: Lincom Europa.

—— and Y.-M. Shin. 2005. The Functional Domain of Concomitance. A Typological Study of Instrumental and Comitative Relations. In *Studia Typologica* 7. 9–104.

Lehrman, A. 2006. Meaning as Grammar. In *Language Sciences* 28. 497–507.

Leibniz, G. W. 1903. *Opuscules et fragments inédits de Leibniz*. Louis Coutrat (ed). Paris: Presses Universitaires de France. [Reprinted 1961. Hildesheim: Georg Olms.]

Lenerz, J. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: G. Narr Verlag.

——1984. *Syntaktischer Wandel und Grammatiktheorie*. Tübingen: Niemeyer.

Leslau, W. 1995. *Reference Grammar of Amharic*. Wiesbaden: Harrassowitz.

Leumann, M., J. B. Hofmann, and A. Szantyr. 1977. *Lateinische Grammatik*. Bd. 1: *Lateinische Laut- und Formenlehre*. Munich: Beck. [Neuausgabe der 1926–1928 erschienenen 5. Aufl. der Lateinischen Grammatik von J. B. Hofmann and A. Szantyr.]

Levelt, W. J. M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: Bradford.

—— A. Roelofs, and A. S. Meyer. 1999. A Theory of Lexical Access in Speech Production. In *Behavioral and Brain Sciences* 22. 1–38.

Levin, B. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago and London: University of Chicago Press.

—— and M. Rappaport Hovav. 1995. *Unaccusativity: At the Syntax–Lexical Semantics Interface*. Cambridge, MA: MIT Press.

—— ——2005. *Argument Realization*. Cambridge: Cambridge University Press.

Levin, J. and D. Massam. 1984. Surface Ergativity: Case/Theta-relations Reexamined. In S. Berman (ed.). *Proceedings of NELS* 15. 286–301.

Levine, R. 1977. *The Skidegate Dialect of Haida*. PhD Dissertation. Columbia University.

Levinson, S. C, S. Meira, and the Language and Cognition Group. 2003. 'Natural Concepts' in the Spatial Topological Domain–Adpositional Meanings in Crosslinguistic Perspective: An Exercise in Semantic Typology. In *Language* 79. 45–516.

Levy, P. 1990. *Totonaco de Papantla, Veracruz*. Mexico: Colegio de México.
Lewis, G. L. 1967. *Turkish Grammar*. Oxford: Clarendon Press.
Leyew, Z. 2003. *The Kemantney Language. A Sociolinguistic and Grammatical Study of Language Replacement*. Cushitic Language Studies 20. Cologne: Rüdiger Köppe.
Li, C. and S. Thompson. 1976. Subject and Topic: A New Typology of Language. In Li, C. (ed.). *Subject and Topic.* New York/San Francisco/London: Academic Press. 457–98.
————1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
————and J. O. Sawyer. 1977. Subject and Word Order in Wappo. In *International Journal of American Linguistics* 43.2. 85–100.
Li, P. J.-K. and S. Tsuchida. 2001. *Pazih Dictionary*. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.
Li, Y.-H. A. 1990. *Order and Constituency in Mandarin Chinese*. Studies in Natural Language & Linguistic Theory 19. Dordrecht: Kluwer Academic Publishers.
Lichtenberk, F. 1983. Relational Classifiers. In *Lingua* 60. 147–76.
Lindenfeld, J. 1973. *Yaqui Syntax*. Berkeley: University of California Press.
Lindsley, J. R. 1975. Producing Simple Utterances: How Far Ahead do we Plan? In *Cognitive Psychology 7*. 1–19.
——1976. Producing Simple Utterances: Details of the Planning Process. In *Journal of Psycholinguistic Research* 5. 331–51.
Lindström, E. 2002. *Topics in the Grammar of Kuot, a Non-Austronesian Language of New Ireland, Papua New Guinea*. PhD thesis, Stockholm University.
Linebarger, M. C., M. F. Schwartz, and E. M. Saffran. 1983. Sensitivity to Grammatical Structure in So-Called Agrammatic Aphasics. In *Cognition* 13. 361–92.
Lomtadze, E. A. 1963. *Ginukhskij dialekt didojskogo jazyka*. Tbilisi.
Loogman, A. 1965. *Swahili Grammar and Syntax*. Louvain: Nauwelaerts.
Lord, C. 1982. The Development of Object Markers in Serial Verb Languages. In S. A. Thompson and P. Hopper (eds.). *Syntax and Semantics* 15. New York: Academic Press. 277–99.
Louwerse, J. 1988. *The Morphosyntax of Una in Relation to Discourse Structure*. Pacific Linguistics, Series B, No. 100. Canberra: Australian National University.
Lukatela, K., S. Crain, and D. Shankweiler. 1988. Sensitivity to Inflectional Morphology in Agrammatism: Investigation of a Highly Inflected Language. In *Brain and Language* 33. 1–15.
Luraghi, S. 1986. On the Distribution of Instrumental and Agent Markers for Human and Non-human Agents of Passive Verbs in Some Indo-European Languages. In *Indogermanische Forschungen* 91. 48–66.
——1987. Patterns of Case Syncretism in Indo-European Languages. In A. G. Ramat, O. Carruba, and G. Bernini (eds.). *Papers from the VIIth International Conference on Historical Linguistics*. Amsterdam: Benjamins. 355–71.
——2000. Synkretismus. In J. Koij, C. Lehmann, and H. Mugdan (eds.). *Handbuch der Morphologie*. Berlin/New York: Mouton de Gruyter.
——2001. Some Remarks on Instrument, Comitative, and Agent in Indo-European. In *Sprachtypologie und Universalienforschung* 54/4. 385–401.
——2003. *On the Meaning of Prepositions and Cases. A Study of the Expression of Semantic Roles in Ancient Greek*. Amsterdam/Philadelphia: Benjamins.

——2004a. The Container Schema in Homeric Greek. In A. Soares de Silva, A. Torres, and M. Gonçalves (eds.). *Linguagem, cultura e cognição: Estudos de Linguística Cognitiva*. Braga: Almedina. 25–41.

——(forthcoming). The History of the Greek Preposition metá: From Polysemy to the Creation of Homonyms. In *Glotta*.

Luria, A. 1976. *Basic Problems of Neurolinguistics*. The Hague: Mouton.

Luutonen, J. 1997. *The Variation of Morpheme Order in Mari Declension*. Helsinki: Suomalais-Ugrilainen Seura.

Lyons, C. 1986. The Syntax of English Genitive Constructions. In *Journal of Linguistics* 22. 123–43.

Lyons, J. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

Lytkin, V. I., K. E. Majtinskaja, and K. Redei (eds.). 1976. *Osnovy finno-ugorskogo jazykoznanija. Marijskij, permskie i ugorskie jazyki*. [The Fundamentals of Finno-Ugric Linguistics. The Mari, Permian and Ugric Languages]. Moscow: Nauka.

MacDonald, L. 1990. *A Grammar of Tauya*. Berlin: Mouton de Gruyter.

MacKenzie, D. N. 1961. *Kurdish Dialect Studies*. London: Oxford University Press.

——1966. *The Dialect of Awroman (Hawrāmān-ī Luhōn). Grammatical Sketch, Texts, and Vocabulary.* Det Kgl. Danske Videnskabernes Selskab., Hist. Filos. Skr. Dan. Vid. Selsk. Vol. 4, 3. Copenhagen.

Mackinnon, R. 1971. *Teach Yourself Gaelic*. London: Hodder and Stoughton.

MacWhinney, B. and J. Osmán-Sági. 1991. Inflectional Marking in Hungarian Aphasics. In *Brain and Language* 41. 165–83.

Maddieson, I. (forthcoming). Areal and Typological Patterns in the Phonology of the Languages of Oceania. In *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society* (2006).

Maffei, A. F. J. 1986. *A Konkani Grammar*. New Delhi: Asian Education Studies.

Magomedbekova, Z. M. 1967. *Akhvakhskij jazyk*. Tbilisi.

——1971. *Karatinskij jazyk*. Tbilisi.

Magomedova, P. T. 1979. Osnovnye voprosy sklonenija v andijskix jazykax v sravnitel'nom aspekte. In U. A. Mejlanova (ed.). *Imennoe sklonenie v dagestanskix jazykax*. Maxačkala: IIJaL. 41–69.

Magometov, A. A. 1963. *Kubachinskij jazyk (issledovanie i teksty)*. Tbilisi.

——1965. *Tabasaranskij jazyk*. Tbilisi: Mecniereba.

——1970. *Agul'skij jazyk*. Tbilisi: Mecniereba.

——1982. *Megebskij dialect darginskogo jazyka (issledovanie i teksty)*. Tbilisi: Mecniereba.

Mahajan, A. 1989. Agreement and Agreement Phrases. In *Functional Heads and Clause Structure*. MIT Working Papers in Linguistics 10. 217–52.

Maizel, S. S. 1957. *Izafet v tureckom jazyke*. Moscow: Izdatel'stvo AN SSSR.

Makhmudova, S. M. 2001. *Morfologija rutul'skogo jazyka* [The Morphology of the Rutul Language]. Moscow: Russian Academy of Sciences.

Malchukov, A. L. 1995. *Even*. Munich: Lincom.

——2000. *Dependency Reversal in Noun–Attribute Constructions: Towards a Typology*. Munich: Lincom Europa.

——2004. *Nominalization/verbalization: Constraining the Typology of Transcategorial Operations*. Munich: Lincom Europa.

MALCHUKOV, A. L. 2005. Case Pattern Splits, Verb Types and Construction Competition. In M. Amberber, and H. de Hoop (eds.). *Competition and Variation in Natural Languages: the Case for Case*. Amsterdam: Elsevier Sciences. 73–119.

—— 2006a. Transitivity Parameters and Transitivity Alternations: Constraining Covariation. In L. Kulikov, A. Malchukov, and P. de Swart (eds.). *Case, Valency and Transitivity*. Amsterdam: Benjamins. 329–58.

—— 2006b. Yakut Interference in North-Tungusic Languages. In H. Boeschoten and L. Johanson (eds.). *Turkic Languages in Contact*. Wiesbaden: Harrassowitz. 122–38.

—— 2007. *'Quirky' case: Rare Phenomena in Case-marking and their Theoretical Implications*. Ms. Max Planck Institute for Evolutionary Anthropology, Leipzig.

—— 2008a. Split Intransitives, Experiencer Objects and 'Transimpersonal' Constructions: (Re-)establishing the Connection. In M. Donohue and S. Wichmann (eds.). *The Typology of Semantic Alignment*. Oxford: Oxford University Press.

—— 2008b. Animacy and Asymmetries in Differential Case Marking. In *Lingua* 118 (Special Issue on Animacy. M. Lamers and P. de Swart eds.). 203–21.

MALÉCOT, A. 1963. Luiseño, a Structural Analysis II: Morphosyntax. In *International Journal of American Linguistics* 29. 196–210.

MALING, J. 1989. Adverbials and Structural Case in Korean. In *Harvard Studies in Korean Linguistics* III. 297–308.

—— 1990. Objects, Adverbs and Grammatical Case in Finnish. In H. Á. Sigurðsson (ed.). *Papers from the Twelfth Scandinavian Conference of Linguistics*. Reykjavík: Linguistic Institute, University of Iceland. 267–78.

—— 1993. Of Nominative and Accusative: The Hierarchical Assignment of Grammatical Case in Finnish. In A. Holmberg, and U. Nikanne (eds.). *Case and Other Functional Categories in Finnish Syntax*. Studies in Generative Grammar 39. Berlin: Mouton de Gruyter. 51–76.

—— 2001. Dative: The Heterogeneity of the Mapping among Morphological Case, Grammatical Functions, and Thematic Roles. In *Lingua* 111. 419–64.

—— 2002. Það rignir þágufalli á Íslandi [It is raining dative in Iceland]: Verbs with Dative Objects in Icelandic. In *Íslenskt mál og almenn málfræði* 24. 31–105. [Also published in a slightly different English version in *Working Papers in Scandinavian Syntax* 70. 1–60. 2002.]

—— and S. KIM. 1992. Case assignment in the Inalienable Possession Construction in Korean. In *Journal of East Asian Linguistics* 1. 37–68.

—— J. JUN, and S. KIM. 2001. Case-marking on Duration Adverbials Revisited. In H. Ahn and S. Kim (eds.). *Selected Papers from the Twelfth International Conference on Korean Linguistics*. Seoul: Kjungjin Publishing Co. 323–35.

MALLINSON, G., and B. J. BLAKE. 1981. *Language Typology: Cross-linguistic Studies in Syntax*. Amsterdam: North Holland.

MANANDISE, E. 1988. *Evidence from Basque for a New Theory of Grammar*. New York/London: Garland Publishing, Inc.

MANASTER-RAMER, A. 1994. The Origin of the Term 'ergative'. In *Sprachtypologie und Universalienforschung* 47.3. 211–14.

MANIACKY, J. 2002. *Tonologie du Ngangela. Variétè de Menongue*. PhD Dissertation. Angola: L'Inalco University.

MANNING, C. D. 1996. *Ergativity: Argument Structure and Grammatical Relations*. Stanford: CSLI.

MARANTZ, A. 1981. *On the Nature of Grammatical Relations*. Ph.D. Dissertation, MIT, Cambridge, MA.

——1984. *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.

——1991. Case and Licensing. In G. Westphal, B. Ao, and H.-R. Chae (eds.). *Proceedings of ESCOL '91*. Cornell Linguistics Club. 234–53.

——1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In A. Dimitriadis and L. Siegel (eds.). *Proceedings of the 21st Annual Penn Linguistics Colloquium, University of Pennsylvania Working Papers in Linguistics*. University of Pennsylvania. 201–25.

MARATSOS, M. P. and M. A. CHALKEY. 1980. The Internal Language of Children's Syntax: The Ontogenesis and Representation of Syntactic Categories. In K. E. Nelson (ed.). *Children's Language*. Vol. 2. New York: Gardner. 127–214.

MARCUS, G. F. 1993. Negative Evidence in Language Acquisition? In *Cognition* 46. 53–85.

MARMO, C. 1994. *Semiotica e linguaggio nella scolastica: Parigi, Bologna, Erfurt*: *1270–1330*: *la semiotica dei Modisti*. Rome: Istituto Storico Italiano per Il Medio Evo.

MARTIN, R. A. 1996. *A Minimalist Theory of PRO and Control*. PhD Dissertation, University of Connecticut, Storrs.

——2001. Null Case and the Distribution of PRO. In *Linguistic Inquiry* 32.1. 141–66.

MARUŠIČ, F. and R. ŽAUCER. In press. On the Intensional FEEL-LIKE Construction in Slovenian: A Case for a Phonologically Null Verb. To appear in: *Natural Language and Linguistic Theory*.

MARX, E. 1999. Gender Processing in Speech Production: Evidence from German Speech Errors. In *Journal of Psycholinguistic Research* 28. 601–21.

MASICA, C. P. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.

MASLOVA, E. 1997. Yukagir Focus System in a Typological Perspective. In *Journal of Pragmatics* 27. 457–75.

——(ed.). 2001. *Yukaghir Texts*. Wiesbaden: Harrassowitz Verlag.

——2003. *A Grammar of Kolyma Yukaghir*. Berlin/New York: Mouton de Gruyter.

——2005. Information Structure in Tundra Yukaghir and Typology of Focus Structures. In J. Fernandez-Vest (ed.). *Les langues ouraliennes aujourd'hui : Approche linguistique et cognitive*. Paris: Honoré Champion. 599–610.

——2006. Information Focus in Relational Clause Structure. In T. Tsunoda and T. Kageyama (eds.). *Voice and Grammatical Relations: In Honor of Masayoshi Shibatani*. Amsterdam/Philadelphia: J. Benjamins. 175–94.

MATEO-TOLEDO, B., V. RODRÍGUEZ, and J. MENDOZA. (in press). La predicación secundaria en awakateko (maya). In R. Zavala and J. Aissen (eds.). *Estudios sobre Predicación Secundaria en Lenguas Mesoamericanas*. Mexico, D.F.: CIESAS.

MATHIASSEN, T. 1996. *A Short Grammar of Lithuanian*. Columbus: Slavica.

MATISOFF, J. A. 1986. Hearts and Minds in Southeast Asian Languages and English: An Essay in the Comparative Lexical Semantics of Psycho-collocations. In *Cahiers de linguistique asie-orientale* 15. 5–57.

MATRAS, Y. 1997. The Typology of Case Relations and Case Layer Distribution in Romani. In Y. Matras, P. Bakker, and H. Kyuchukov (eds.). *The Typology and Dialectology of Romani*. Amsterdam: Benjamins. 61–93.

MATTHEWS, P. H. 1972. *Inflectional Morphology: A Theoretical Study based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.

MATTHEWS, P. H. 1991. *Morphology: an Introduction to the Theory of Word-structure*, 2nd ed. Cambridge: Cambridge University Press.

MAYNARD. S. K. 1980. *Discourse Functions of the Japanese Theme Marker 'wa'*. PhD dissertation, Northwestern University.

MAZUKA, R. and N. NAGAI (eds.). 1995. *Japanese Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum.

MCCAWLEY, N. 1976. From OE/ME 'Impersonal' to 'Personal' Constructions: What is a 'Subject-Less' S?. In *Papers from the Parasession on Diachronic Syntax*. Chicago: Chicago Linguistic Society. 192–204.

MCFARLAND, C. 1978. *Definite Object and Subject Selection in Philippine Languages*. Studies in Philippine Linguistics 2 (1).

MCGREGOR, W. 1998. Optional Case-Marking in Gooniyandi Revisited: Implications to a Theory of Marking. In *Leuvense Bijdragen* 87. 491–534.

—— (to appear). Indexicals as Sources of Case Markers. In F. Josephson, and I. Söhrman (eds.). *Forms and Functions: Aspect, Tense, Mood, Diathesis and Valency*. Proceedings of the First Colloquium on Language Typology in a Diachronical Perspective held at Göteborg University, November 2004. Amsterdam: Benjamins. 19–21.

MCGUCKIN, C. 2002. Gapapaiwa. In J. Lynch, M. Ross, and T. Crowley (eds.). *The Oceanic Languages*. Richmond (Surrey): Curzon Press. 297–320.

MCLENDON, S. 1975. *A Grammar of Eastern Pomo*. Berkeley: University of California Press.

——1978. Ergativity, Case and Transitivity in Eastern Pomo. In *International Journal of American Lingusitics* 44. 1–9.

MEIER, H. 1967. *Deutsche Sprachstatistik*. Vols. 1 and 2. Hildesheim: Olms.

MEIKLEJOHN, J. M. D. 1892. *A New Grammar of the English Tongue*, 9th edn. London: Simpkin, Marshall, Hamilton, Kent, & Co.

MEILLET, A. 1936. *Esquisse d'une grammaire comparée de l'arménien classique*, 2nd edn. Vienna: Imprimerie des Frères Mekhitaristes.

——1949. *Caractères generaux des langues germaniques*. Paris: Hachette.

——1950. *Les dialects indo-européens*. Paris: Champion.

MEIRA, S. 1999. *A Grammar of Tiriyó*. Unpublished PhD Dissertation at Rice University.

MEL'ČUK, I. 1986. Toward a Definition of Case. In R. D. Brecht and J. S. Levine (eds.). *Case in Slavic*. Columbus, OH: Slavica Publishers, Inc. 35–85.

——1988. Is there an Ergative Construction in Lezgian? In I. A. Mel'čuk. *Dependency Syntax: Theory and Practice*. Albany: State University of New York Press. 207–49.

——1998. *Kurs obščej morfologii*, vol.2. Moscow: Jazyki russkoj kul'tury.

MELINGER, A. and A. CLELAND. 2005. *Noun Phrase Structure Priming within a Sentence: The Role of Grammatical Function and Linear Order*. Paper presented at the 18th Annual CUNY Conference on Human Language Processing, Tucson, Arizona.

—— —— (in preparation). *Noun Phrase Structural Priming at the Beginnings and Ends of Sentences*.

—— and C. DOBEL. 2005. Lexically-Driven Syntactic Priming. In *Cognition* 98. B11–B20.

MENG, M. and M. BADER. 2000. Ungrammaticality Detection and Garden-Path Strength: Evidence for Serial Parsing. In *Language and Cognitive Processes* 15. 615–66.

MENN, L. and B. MACWHINNEY. 1984. The Repeated Morph Constraint: Toward an Explanation. In *Language* 60. 519–41.

—— and L. K. OBLER. 1990. *Agrammatic Aphasia: a Cross-Language Narrative Sourcebook*. Amsterdam: John Benjamins.

——M. O'Connor, L. K. Obler, and A. Holland. 1995. *Non-Fluent Aphasia in a Multilingual World*. Amsterdam: John Benjamins.

Mennecier, P. 1995. *Le tunumiisut, dialecte inuit du Groenland oriental: description et analyse*. Paris: Klincksieck.

Menovshikov, G. A. and N. B. Vakhtin. 1990. *Eskimosskij jazyk* [The Eskimo Language]. Leningrad: Prosveščenie.

Merdanova, S. R. 2004. *Morfologija i grammatičeskaja semantika agul'skogo jazyka*. Moscow: Institut jazykoznanija RAN.

Merlan, F. 1982. *Mangarayi*. Lingua Descriptive Series 4. Amsterdam, North-Holland: Elsevier.

——1985. Split Intransitivity: Functional Oppositions in Intransitive Inflections. In J. Nichols and A. C. Woodbury (eds.). *Grammar Inside and Outside the Clause: Some Approaches to Theory from the Field*. Cambridge: Cambridge University Press. 324–62.

——1989. *Mangarayi*. London: Routledge.

Michaelis, S. and M. Rosalie. 2000. Polysémie et cartes sémantiques: le relateur (av)ek en créole seychellois. *Études Créoles* XXIII/2. 79–100.

Miller, A. 2001. *A Grammar of Jamul Tiipay*. Mouton Grammar Library 23. Berlin/New York: Mouton de Gruyter.

Miller, B. V. 1953. *Talyšskij jazyk*. Moscow: Akademija Nauk.

Miller, M. 1999. *Desano Grammar. Studies in the Languages of Colombia*. The University of Texas at Arlington: Summer Institute of Linguistics.

Mills, A. E. 1985. The Acquisition of German. In D. I. Slobin (ed.). *The Crosslinguistic Study of Language Acquisition*. Vol. 1. Hillsdale, NJ: Erlbaum. 141–254.

Miloradović, S. 2003. *Upotreba padežnih oblika u govoru Paraćinskog Pomoravlja. Balkanistički i etnomigracioni aspekt*. Belgrade: Etnografski institut.

——and M. Greenberg. 2001. The Transition from South Slavic to Balkan Slavic: Key Morphological Features in Serbian Transitional Dialects. In V. A. Friedman, and D. L. Dyer (eds.). *Of All the Slavs my Favorites. In Honor of Howard I. Aronson*. Indiana Slavic Studies, Vol. 12. Bloomington, IN: Slavica. 311–22.

Mišeska-Tomić, O. (ed.). 2004. *Balkan Syntax and Semantics*. Amsterdam/Philadelphia: John Benjamins.

Mitchell, D. C. 1994. Sentence Parsing. In M. A. Gernsbacher (ed.). *Handbook of Psycholinguistics*. San Diego: Academic Press. 375–409.

Mithun, M. 1984. The Evolution of Noun Incorporation. In *Language* 60. 847–94.

——1985. Disagreement: The Case of Pronominal Affixes and Nouns. In D. Tannen and J. E. Alatis (eds.). *Georgetown Round Table on Languages and Linguistics 1985: Languages and Linguistics: The Interdependence of Theory, Data and Application*. Washington, DC: Georgetown University Press. 50–66.

——1991. Active/Agentive Case Marking and its Motivations. In *Language* 67.3. 510–46.

——1999. *The Languages of Native North America*. Cambridge: Cambridge University Press.

——(2008). The Emergence of Agent/Patient Systems. In M. Donohue and S. Wichmann (eds.). *The Typology of Semantic Alignment*. Oxford: Oxford University Press.

Mladenov, M. and K. Steinke. 1978. Die Ergebnisse der neueren bulgarischen Dialektforschung im Lichte der Balkanologie. In *Zeitschrift für Balkanologie*. Vol. XIV. Wiesbaden. 68–82.

Mohanan, T. 1994. *Argument Structure in Hindi*. Stanford: CSLI Publications.

Moignet, G. 1976. *Grammaire de l'ancien français. Morphologie–Syntaxe*. 2nd edn. rev. and corr. Paris: Klincksieck.

Montague, R. 1973. The Proper Treatment of Quantification in Ordinary English. In J. Hintikka, J. Moravcsik, and P. Suppes (eds.). *Approaches to Natural Language*. Dordrecht: Reidel. 221–42.

Moravcsik, E. A. 1974. Object–Verb Agreement. In *Working Papers in Language Universals* 15. Stanford: Stanford University Press. 25–40.

——1978. On the Case Marking of Objects. In J. Greenberg (ed.). *Universals of Human Language*. Vol. 4. Stanford: Stanford University Press. 249–89.

——1995. Summing up Suffixaufnahme. In F. Plank (ed.). *Double Case: Agreement by Suffixaufnahme*. Oxford: OUP. 451–84.

——2003. Inflectional Morphology in the Hungarian Noun Phrase, in F. Plank (ed.). *Noun Phrase Structure in the Languages of Europe*. Berlin: Mouton de Gruyter. 397–466.

Morgenstierne, G. 1949. The Language of the Prasun Kafirs. In *Norsk Tidsskrift for sprogvidenskap* 15. 188–334.

Morikawa, H. 1989. *Acquisition of Case Marking and Predicate–Argument Structures in Japanese: A Longitudinal Study of Language Acquisition Mechanisms*. Dissertation. Lawrence, KS: University of Kansas.

Morimoto, Y. 2008. From Topic to Subject Markers: Implications for a Typology of Subject Marking. In H. de Hoop and P. de Swart (eds.). *Differential Subject Marking*. Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer. 199–223.

——and P. de Swart. 2006. Language Variation and Historical Change: The Spread of DOM in Spanish. In J. Doetjes and P. González (eds.). *Selected Papers from Going Romance 2004*. Amsterdam/Philadelphia: John Benjamins. 225–45.

Morin, Y.-C. and E. Tiffou. 1988. Passives in Burushaski. In M. Shibatani (ed.). 1988. *Passive and Voice*. Amsterdam: John Benjamins. 493–524.

Morse-Gagne, E. E. 2003. *Viking Pronouns in England: Charting the course of THEY, THEIR, and THEM*. Dissertation, University of Pennsylvania.

Mosel, U. 1987. Subject in Samoan. In D. C. Laycock, and W. Winter (eds.). *A World of Language*. Canberra: Pacific Linguistics. 455–79.

——and E. Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.

Moshinsky, J. 1974. *A Grammar of South Eastern Pomo*. Berkeley: University of California Publications in Linguistics.

Mous, M. 1993. *A Grammar of Iraqw*. Hamburg: Buske.

Mullen, D. S. 1986. *Issues in the Morphology and Phonology of Amharic: The Lexical Generation of Pronominal Clitics*. Unpublished doctoral dissertation, University of Ottawa, Ottawa.

Müller, G., L. Gunkel, and G. Zifonun (eds.). 2004. *Explorations in Nominal Inflection*. Berlin: Mouton de Gruyter.

Muravyova, I. A. 1998. Chukchee (Paleo-Siberian). In A. Spencer and A. Zwicky (eds.). *Handbook of Morphology*. Oxford: Blackwell Publishers. 521–38.

Murkelinskij, G. B. 1971. *Grammatika lakskogo jazyka*. Maxačkala.

Mutalov, R. O. 2002. *Glagol darginskogo jazyka*. Maxačkala.

Næss, Å. 2004a. *Transitivity: From Semantics to Structure*. Doctoral dissertation. University of Nijmegen.

——2004b. What Markedness Marks: The Markedness Problem with Direct Objects. In *Lingua* 114. 1186–1212.
——2006. Case Semantics and the Agent–Patient Opposition. In L. Kulikov, A. Malchukov, and P. de Swart (eds.). *Case, Valency and Transitivity*. Amsterdam and Philadelphia: John Benjamins. 309–28.
Nagano, Y. 1984. *A Historical Study of the Gyarong Verb System*. Tokyo: Seishido.
Nagaraja, K. 1985. *Khasi: A Descriptive Analysis*. Pune: Deccan College Post-Graduate and Research Institute.
——1999. *Korku Language: Grammar, Texts, and Vocabulary*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
Nakamura, W. 1999. Functional Optimality Theory: Evidence from Split Case Systems. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan, and K. Wheatley (eds.). *Functionalism and Formalism in Linguistics*. Vol. 2. Amsterdam: Benjamins. 253–77.
Narasimhan, B. 2005. Splitting the Notion of 'Agent': Case-marking in Early Child Hindi. *Journal of Child Language* 32. 787–803
——N. Budwig, and L. Murty. 2005. Argument Realization in Hindi Caregiver–Child Discourse. In *Journal of Pragmatics* 37:4. 461–95.
Narrog, H. and S. Ito. (2007). Reconstructing Semantic Maps – the Comitative–Instrumental Area. In *Sprachtypologie und Universalienforschung* 60, 4. 273–92.
Nedjalkov, I. 1997. *Evenki*. London: Routledge.
Neeleman, A. 1994. *Complex Predicates*. PhD dissertation, Utrecht University.
——and F. Weerman. 1999. *Flexible Syntax: A Theory of Case and Arguments*. Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer Academics.
——and H. van de Koot. 2006. *Theta Theory*. Ms. UCL.
—— —— 2007. *The Nature of Discourse Templates*. Ms. UCL.
Neidle, C. 1988. *The Role of Case in Russian Syntax*. Dordrecht: Kluwer.
Nettle, D. 1999. *Linguistic Diversity*. Oxford: Oxford University Press.
Neu, E. 1980. *Studien zum endungslosen Lokativ des Hethitischen*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
Nevis, J. A. 1988. A Morphotactic Paradox in Northern Saame: Comitative *guim*. In *Ural-Altaische Jahrbücher* 54 (8). 38–50.
Newman, J. 1988. Recipients and 'give' Constructions. In W. Van Langendonck and W. Van Belle (eds.). *The Dative.* Volume 2: *Theoretical and Contrastive Studies*. Amsterdam/Philadelphia: John Benjamins. 1–28.
van Nice, K. Y. and R. Dietrich. 2003. Task-sensitivity of Animacy Effects: Evidence from German Picture Descriptions. In *Linguistics* 5. 825–49.
Nichols, J. 1986. Head-Marking and Dependent-Marking Grammar. In *Language* 62. 56–119.
——1988. On Alienable and Inalienable Possession. In W. Shipley (ed.). *In Honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics.* Berlin: Mouton de Gruyter. 557–609.
——1992. *Linguistic Diversity in Space and Time*. Chicago: The University of Chicago Press.
——1994a. Chechen. In R. Smeets (ed). *Indigenous Languages of the Caucasus: North East Caucasian Languages*. Delmar, NY: Caravan Books. 1–78.
——1994b. Ingush. In R. Smeets (ed.). *Indigenous Languages of the Caucasus: North East Caucasian Languages.* Delmar, NY: Caravan Books. 79–147.

Nichols, J. 1997a. Sprung from Two Common Sources: Sahul as a Linguistic Area. In P. McConvell (ed.). *Archeology and Linguistics: Global Perspectives on Ancient Australia*. Melbourne.

—— 1997b. Modeling Ancient Population Structures and Population Movement in Linguistics and Archeology. In *Annual Review of Anthropology* 26. 359–84.

—— (forthcoming). A Case of Rare Fluid Intransitivity in Eurasia: Russian. In *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society* (2006).

—— 2007. Why are Stative–active Languages Rare in Eurasia? Typological Perspective on Split Subject Marking. In M. Donohue and S. Wichmann (eds.). *The Typology of Semantic Alignment Systems*. Oxford: Oxford University Press. 121–39.

Nikanne, U. 1990. *Zones and Tiers*. Helsinki: Finnish Literature Society.

Nikiforidou, K. 1991. The Meaning of the Genitive. In *Cognitive Linguistics* 2/2. 149–205.

Nikolaeva, I. 1999. *Ostyak*. Munich: Lincom.

—— and M. Tolstaya. 2001. *A Grammar of Udihe*. Berlin: Mouton de Gruyter.

Nilsen, D. F. 1973. *The Instrumental Case in English: Syntactic and Semantic Considerations*. The Hague: Mouton.

Noda, M. 1996. *'Wa' to 'ga'* [The particles ∼*wa* and ∼*ga*]. Tokyo: Kuroshio Publishing.

Nolasco, R. 2005. What Ergativity in Philippine Languages Really Means. In *The Proceedings of the Taiwan–Japan Joint Workshop on Austronesian Languages*. The Humanities Research Center, National Science Research Council. Taipei. 215–61.

Nomura, M. 2005. *Nominative Case and AGREE(ment)*. PhD Dissertation, University of Connecticut, Storrs.

Noonan, M. P. 2003. Chantyal. In G. Thurgood and R. J. LaPolla (eds.). *The Sino-Tibetan Languages*. London: Routledge. 315–35.

Nooteboom, S. G., J. P. L. Brokx, and J. J. de Rooij. 1978. Contributions of Prosody to Speech Perception. In W. J. M. Levelt and G. B. Flores d'Arcais (eds.). *Studies in the Perception of Language*. New York: Wiley. 75–107.

Nordlinger, R. 1998a. *A Grammar of Wambaya, Northern Territory (Australia)*. Canberra: Pacific Linguistics

—— 1998b. *Constructive Case: Evidence from Australian Languages*. Stanford: CSLI Publications.

—— 2000. Australian Case Systems: Towards a Constructive Solution. In M. Butt and T. H. King (eds.). *Argument Realization*. Stanford, CA: CSLI Publications. 41–72.

Ochs, E. and B. B. Schieffelin. 1995. The Impact of Language Socialization on Grammatical Development. In P. Fletcher and B. MacWhinney (eds.). *The Handbook of Child Language*. Oxford: Blackwell. 73–95.

Ogawa, A. 1997. Argumenterhöhung im Sprachvergleich. Dativ im Deutschen vs. Postposition '-ni', Multiple-Subjekt und Adversativ-Passiv im Japanischen. In *Sprachwissenschaft* 22. 181–206.

—— 2006. Meteorological and Chronological Expressions in Japanese and some Other Languages. In *Studia Universitatis Babeș-Bolyai* 1, Cluj/Romania. 33–45.

Olson, C. 1992. Gumawana (Amphlett Islands, Papua New Guinea): Grammar Sketch and Texts. In M. D. Ross (ed.). *Papers in Austronesian Linguistics* 2. Canberra: Australian National University. 251–430.

Onishi, M. 1994. Semantic Primitives in Japanese. In C. Goddard and A. Wierzbicka (eds.). S*emantic and Lexical Universals: Theory and Empirical Findings.* Amsterdam/Philadelphia. John Benjamins. 361–86.

——1997. The Grammar of Mental Predicates in Japanese. In *Language Sciences* 19(3). 219–33.

——2001. Introduction: Non-Canonically Marked Subjects and Objects: Parameters and Properties. In A. Y. Aikhenvald, R. M. W. Dixon, and M. Onishi (eds.). *Non-Canonical Marking of Subjects and Objects*. Amsterdam: John Benjamins. 1–51.

Onoe, K. 1995. 'Wa' no imi bunka no ronri [On the process of polysemy of ~*wa*]. In *Gengo* (Language) 24/11. 28–37.

Orr, R. 2000. *Common Slavic Nominal Morphology: A New Synthesis.* Bloomington: Slavica.

Otoguro, R. 2006. *Morphosyntax of Case. A Theoretical Investigation of the Concept*. PhD dissertation, University of Essex.

Padley, G. A. 1976. *Grammatical Theory in Western Europe 1500–1700: The Latin Tradition*. Cambridge: Cambridge University Press.

Palancar, E. L. 1999. Instrumental Prefixes in Amerindian Languages: An Overview to their Meanings, Origin, and Functions. In *Sprachtypologie und Universalienforschung* 52/2. 151–66.

——2001. Emergent Markers: The Relation of Instruments, Causes, and Agents in Grammatical Constructions. In *Sprachtypologie und Universalienforschung* 54/4. 365–84.

——2002. *The Origin of Agent Markers*. Studia Typologica 5. Berlin: Akademie Verlag.

Palmer, F. R. 1994. *Grammatical Roles and Relations*. Cambridge: Cambridge University Press.

Panfilov, V. Z. 1962. *Grammatika nivxskogo jazyka*. [A Grammar of Nivkh]. Moscow: Nauka.

——1965. *Grammatika nivxskogo jazyka*, 2. Leningrad: Nauka.

Panov, M. V. 1997. *Sovremennyj russkij jazyk. Fonetika*. Moscow: Vyšaja Škola.

Pappert, S., J. Schliesser, D. P. Janssen, and T. Pechmann. 2007. Corpus- and Psycholinguistic Investigations of Linguistic Constraints on German Object Order. In A. Spaeth (ed.). *Interfaces and Interface Conditions*. Berlin: DeGruyter.

Park, K.-S. 1995. *The Semantics and Pragmatics of Case Marking in Korean: A Role and Reference Grammar Account*. Unpublished Ph.D. dissertation, University at Buffalo. Available on RRG web site: http://linguistics.buffalo.edu/research/rrg.html

Parsons, T. 1990. *Events in the Semantics of English*. Cambridge, MA: MIT Press.

——1995. Thematic Relations and Arguments. In *Linguistic Inquiry* 26. 635–62.

Partee, B. H. 1986. Noun Phrase Interpretation and Type-Shifting Principles. In J. Groenendijk, D. de Jongh, and M. Stokhof (eds.). *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Dordrecht: Foris. 115–43.

——1997. Genitives – A Case Study. Appendix to T. M. V. Janssen. 'Compositionality'. In J. van Benthem and A. ter Meulen (eds.). *The Handbook of Logic and Language*. Amsterdam: Elsevier. 464–70.

——and V. Borschev. 2000. Possessives, *favourite* and Coercion. In A. Riehl and R. Daly (eds.). *Proceedings of ESCOL99*. Ithaca, NY: CLC Publications. 173–90.

———2004. The Semantics of Russian Genitive of Negation: The Nature and Role of Perspectival Structure. In R. B. Young (ed.). *Proceedings from SALT XIV*. Ithaca: CLC Publications. 212–34.

Paul, H. 1920. *Prinzipien der Sprachgeschichte*. 5th edn. Tübingen: Max Niemayer Verlag.

——P. Wiehl, and S. Grosse. 1989. *Mittelhochdeutsche Grammatik* (23rd edition). Tübingen: Max Niemeyer.

Paul, L. 1998. *Zazaki*. Wiesbaden: Reichert.
—— 2002. *Grammatical and Philological Studies on the Early Judaeo-Persian Texts from the Cairo Geniza*. Habilitationsschrift, Univ. Goettingen.
Payne, D. L. and I. Barshi (eds.). 1999. *External Possession*. Amsterdam: John Benjamins.
Payne, J. 1980. The Decay of Ergativity in Pamir Languages. In *Lingua* 51. 147–86.
—— 1989. Pāmir Languages. In R. Schmitt (ed.). *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert. 417–44.
Pearson, M. 2005. The Malagasy Subject/Topic as an A′-element. In *Natural Language and Linguistic Theory* 23. 381–457.
Pederson, E. 1995. Language as Context, Language as Means: Spatial Cognition and Habitual Language Use. In *Cognitive Linguistics* 6 (1). 33–62.
—— E. Danziger, D. Wilkins, S. Levinson, S. Kita, and G. Senft. 1998. Semantic Typology and Spatial Conceptualization. In *Language* 74. 557–89.
Pedersen, H. 1905. Zur armenischen Sprachgeschichte. In *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen (Kuhns Zeitschrift)* 38: 194–240. [Reprinted in H. Pedersen. 1982. *Kleine Schriften zum Armenischen*. Hildesheim: Olms. 56–240.]
Peeters, B. (ed.). 2006. *Semantic Primes and Universal Grammar: Empirical Evidence from the Romance Languages.* Amsterdam: John Benjamins.
Penny, R. 2002. *A History of the Spanish Language*. 2nd edition. Cambridge: Cambridge University Press.
Pereltsvaig, A. 1999. Cognate Objects in Russian: Is the Notion 'Cognate' Relevant for Syntax? In *Canadian Journal of Linguistics* 44. 267–91.
Perlmutter, D. M. 1978. Impersonal Passives and the Unaccusative Hypothesis. In *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*. 157–89.
—— and J. Moore. 2002. Language-internal Explanation: The Distribution of Russian Impersonals. In *Language* 78. 619–50.
—— and P. M. Postal. 1983a. Some Proposed Laws of Basic Clause Structure. In D. M. Perlmutter (ed.). *Studies in Relational Grammar 1.* University of Chicago Press, Chicago. 81–128.
—— —— 1983b. Toward a Universal Characterization of Passivization. In D. M. Perlmutter (ed.). *Studies in Relational Grammar 1.* Chicago: The University of Chicago Press. 3–29.
—— 1984. The 1-advancement Exclusiveness Law. In D. Perlmutter and C. Rosen (eds.). *Studies in Relational Grammar 2*. Chicago: University of Chicago Press. 81–125.
Pesetsky, D. 1992. *Zero Syntax II: An Essay on Infinitives.* Unpublished manuscript, MIT, Cambridge, MA.
—— 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge, MA: The MIT Press.
—— and E. Torrego. 2001. T-to-C Movement: Causes and Consequences. In M. Kenstowicz (ed.). *Ken Hale: a Life in Language*. Cambridge, MA: MIT Press. 355–426.
—— —— 2004. Tense, Case, and the Nature of Syntactic Categories. In Jacqueline Guéron and Jacqueline Lecarme (eds.). *The Syntax of Time*. Cambridge, MA: MIT Press. 495–537.
Peters, A. M. and L. Menn. 1993. False Starts and Filler Syllables: Ways to Learn Grammatical Morphemes. In *Language* 69. 742–77.
Peterson, J. 1998. *Grammatical Relations in Pali and the Emergence of Ergativity in Indo-Aryan*. Munich: Lincom Europa.

PETERSON, R. R. and P. SAVOY. 1998. Lexical Selection and Phonological Encoding during Language Production: Evidence for Cascaded Processing. In *Journal of Experimental Psychology: Learning, Memory and Cognition* 24. 539–57.

PETERSON, T. 2006. Issues of Morphological Ergativity in the Tsimshian Languages: Agreement, Determiners and the Reconstruction of Case. In L. Kulikov, A. Malchukov, and P. de Swart (eds.). *Case, Valency and Transitivity*. Amsterdam and Philadelphia: John Benjamins. 65–91.

PICKERING, M. 1999. Sentence Comprehension. In S. Garrod and M. Pickering (eds.). *Language Processing*. Hove: Psychology Press. 123–53.

PICKERING, M. J. and H. P. BRANIGAN. 1998. The Representation of Verbs: Evidence from Syntactic Priming in Language Production. In *Journal of Memory and Language* 39. 633–51.

——— ———1999. Syntactic Priming in Language Production. In *Trends in Cognitive Sciences* 3. 136–41.

——— ———and J. F. MCLEAN. 2002. Constituent Structure is Formulated in One Stage. *Journal of Memory and Language* 46. 586–605.

PILHOFER, G. 1933. *Grammatik der Kâte-Sprache in Neuguinea*. Berlin: Reimer.

PIÑANGO, M. M. 2002. Cortical Reflections of Two Pronominal Relations. In H. J. Simon and H. Wiese (eds.). *Pronouns – Grammar and Representation*. Amsterdam: John Benjamins. 233–52.

——— and P. BURKHARDT. 2005. Pronominal Interpretation and the Syntax-Discourse Interface: Real-time Comprehension and Neurological Properties. In A. Branco, T. McEnery, and R. Mitkov (eds.). *Anaphora Processing: Linguistic, Cognitive and Computational Modelling*. Amsterdam/Philadelphia: John Benjamins.

PINAULT, G.-J. Introduction au tokharien. In *LALIES* 7. 1–224.

PINKER, S. 1981. A Theory of Graph Comprehension. In R. Freedle (ed.). *Artificial Intelligence and the Future of Testing*. Hillsdale, NJ: Erlbaum.

———1984. *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

———1989. *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

PINKSTER, H. 1990. *Latin Syntax and Semantics*. London: Routledge.

PISCHEL, R. 1981. *A Grammar of the Prakrit Languages*. Delhi: Banarsidass.

PITKIN, H. 1985. *Wintu Dictionary*. Berkeley: University of California Press.

PITMAN, D. 1980. *Bosquejo de la Gramatica Araona*. Notas Lingüísticas 9. Riberalta, Bolivia: Instituto Lingüístico de Verano.

PLANK, F. 1979. The Functional Basis of Case Systems and Declension Classes: From Latin to Old French. In *Linguistics* 17(7/8). 611–40.

———1985. The Extended Accusative/Restricted Nominative in Perspective. In F. Plank (ed.). *Relational Typology*. Berlin: Mouton. 269–311.

——— (ed.). 1991a. *Paradigms: The Economy of Inflection*. Berlin: Mouton.

———1991b. Rasmus Rask's dilemma. In F. Plank (ed.). *Paradigms: The Economy of Inflection*. Empirical Approaches to Language Typology 9. Berlin/New York: Mouton de Gruyter. 161–96.

——— (ed.). 1995a. *Double Case. Agreement by Suffixaufnahme*. Oxford: Oxford University Press.

PLANK, F. 1995b. (Re-)Introducing Suffixaufnahme. In F. Plank (ed.). *Double Case. Agreement by Suffixaufnahme*. Oxford: Oxford University Press. 3–110.

—— (ed.). 2003. *Noun Phrase Structure in the Languages of Europe*. Berlin: Mouton de Gruyter.

PLUNGIAN, V. 1995. *Dogon*. Munich: Lincom Europa.

—— 2002. K semantike russkogo lokativa ('vtorogo predložnogo' padeža) [Towards the semantics of the second Russian locative ('second prepositional' case)]. In *Semiotika i informatika* 37. 229–54.

POLLARD, C. and I. A. SAG. 1994. *Head-Driven Phrase Structure Grammar*. Chicago and London: The University of Chicago Press.

POPE, M. K. 1934. *From Latin to Modern French with Especial Consideration of Anglo-Norman. Phonology and Morphology*. Manchester: Manchester University Press.

POPKONSTANTINOV, K. and O. KRONSTEINER. 1997. *Altbulgarische Inschriften*. vol. 2. Salzburg: Die slawischen Sprachen, Bd. 52.

POPOV, B. 1984. Položaj srpskohrvatskog jezika u balkanskom jezičkom savezu. In *Južnoslovenski filolog*. Knj. XL. Belgrade. 21–43.

POTTER, M. C. and L. LOMBARDI. 1990. Regeneration in the Short-term Recall of Sentences. In *Journal of Memory and Language* 29. 633–54.

POUDEL, K. P. 2006. *Dhankute Tamang grammar*. Munich: Lincom Europa.

PRIMUS, B. 1999. *Cases and Thematic Roles – Ergative, Accusative and Active.* Tübingen: Niemeyer.

—— 2006a. Mismatches in Semantic Role Hierarchies and the Dimensions of Role Semantics. In I. Bornkessel et al. (eds.). *Semantic Role Universals and Argument Linking: Theoretical, Typological and Psycholinguistic Perspectives*. Berlin: Mouton de Gruyter. 53–87.

—— 2006b. Case Marking Typology. In J. J. Song. (ed.). *The Handbook of Linguistic Typology*. Oxford: Oxford University Press.

PRINCE, A. and P. SMOLENSKY. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.

PRZEPIÓRKOWSKI, A. (1999). On Complements and Adjuncts in Polish. In R. Borsley and A. Przepiórkowski (eds.). *Slavic in Head-Driven Phrase Structure Grammar*. Stanford: CSLI Publications. 183–210.

PULKINA, I. and E. ZAKHAVA-NEKRASOVA. 1997. *Russian. A Practical Grammar With Exercises*. 7th edition. Moscow: Russky Yazyk Publishers.

PULLUM, G. K. 1988. Citation Etiquette beyond Thunderdome. In *Natural Language and Linguistic Theory* 6: 579–588. [Reprinted in G. Pullum. 1991. *The Great Eskimo Vocabulary Hoax and other Irreverent Essays on the Study of Language*. Chicago: The University of Chicago Press.]

PYE, C. L. 1990. The Acquisition of Ergative Languages. In *Linguistics* 28. 1291–330.

QUIRK, R., S. GREENBAUM, G. LEECH, and J. SVARTVIK. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

R Development Core Team. 2006. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing (www.r-project.org).

RADFORD, A. 1990. *Syntactic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammars of English*. Oxford: Blackwell.

RAMCHAND, G. 1993. *Aspect and Argument Structure in Modern Scottish Gaelic*. PhD dissertation, Stanford University.

——1997. *Aspect and Predication: The Semantics of Argument Structure*. Oxford: Clarendon Press.

——2007. *A First Phase Syntax*. Oxford: Oxford University Press.

Randal, S. 1998. A Grammatical Sketch of Tennet. In G. J. Dimmendaal and M. Last (eds.). *Surmic Languages and Cultures.* Nilo-Saharan 13. Cologne: Rüdiger Köppe. 219–72.

——2000. Tennet's Ergative Origins. In *Occasional Papers in the Study of Sudanese Languages*, 8. Nairobi: Summer Institute of Linguistics. 67–80.

Rapoport, T. 1991. Adjunct-predicate Licensing and D-structure. In S. Rothstein (ed.). *Perspectives on Phrase Structure: Heads and Licensing* (Syntax and Semantics 25). San Diego, CA: Academic Press. 159–87.

Rappaport Hovav, M. and B. Levin. 2005. All Dative Verbs are not Created Equal. Unpublished ms. Jerusalem and Stanford, CA: The Hebrew University of Jerusalem and Stanford University. [http://www.stanford.edu/∼bclevin/dat05.pdf]

Rappaport, M. 1983. On the Nature of Derived Nominals. In L. Levin, M. Rappaport, and A. Zaenen (eds.). *Papers in Lexical-Functional Grammar*. Indiana University Linguistics Club. [Reprinted in M. Butt and T. H. King (eds.). *Lexical Semantics in LFG*. Stanford, CA: CSLI Publications, 2006.]

Ratitamkul, T., A. E. Goldberg, and C. Fisher. 2004. *The Role of Discourse Context in Determining the Argument Structure of Novel Verbs with Omitted Arguments.* Presentation at the Child Language Research Forum, Stanford University, April 2004. Online Proceedings: http://cslipublications.stanford.edu/CLRF/2004/04-Ratitamkul-etal.pdf

van Reenen, P. and L. Schøsler. 2000. Declension in Old and Middle French: Two Opposing Tendencies. In J. C. Smith and D. Bentley (eds.). *Historical Linguistics 1995.* Vol. I: *General Issues and Non-Germanic Languages*. Amsterdam: Benjamins. 327–44.

Reesink, G. 1987. *Structures and their Functions in Usan: a Papuan Language of Papua New Guinea*. Amsterdam/Philadelphia: John Benjamins.

Refsing, K. 1986. *The Ainu Language*. Aarhus: Aarhus University Press.

Reh, M. 1986. Where have All the Prefixes Gone? In *Afrikanistische Arbeitspapiere* 5. 121–34.

——1996. *Anywa Language: Description and Internal Reconstructions*. Nilo-Saharan 11. Cologne: Rüdiger Köppe.

Renck, G. L. 1975. *A Grammar of Yagaria*. Pacific Linguistics B-40. Canberra: Australian National University.

Rhys Jones, T. J. 1977. *Living Welsh*. London: Hodder and Stoughton.

van Riemsdijk, H. 1983. The Case of German Adjectives. In F. Heny and T. Richards (eds.). *Auxiliaries and Related Puzzles.* Dordrecht: Reidel. 223–52.

—— and R. Huybregts. 2001. Location and Locality. In M. van Oostendorp and E. Anagnostopoulou (eds.). *Progress in Grammar*. Amsterdam: Meertens. 1–23.

Rijkhoff, J. and D. Bakker. 1998. Language Sampling. In *Linguistic Typology* 2. 263–314.

Rispoli, M. 1991. The Acquisition of Verb Subcategorization in a Functionalist Framework. In *First Language* 11. 41–63

——1999. Case and Agreement in English Language Development. In *Journal of Child Language* 26. 357–72.

Ritter, E. and S. T. Rosen. 2001. The Interpretive Value of Object Splits. In *Language Sciences* 23. 425–51.

—— ——2005. Topic or Aspect. In P. Kempchinksy and R. Slabakova (eds.). *Aspectual Inquiries*. Dordrecht: Springer. 21–40.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.
——1997. The Fine Structure of the Left Periphery. In L. Haegeman (ed.). *Elements of Grammar*. Dordrecht: Kluwer. 281–337.
Robins, R. H. 1993. *The Byzantine Grammarians: Their Place in History*. Berlin, New York: Mouton de Gruyter.
——1997. *A Short History of Linguistics*, 4th edn. London: Longman.
Roelofs, A. 1992. A Spreading-Activation Theory of Lemma Retrieval in Speaking. In *Cognition* 42. 107–42.
——2000. Weaver++ and Other Computational Models of Lemma Retrieval and Word-form Encoding. In L. Wheeldon (ed.). *Aspects of Language Production*. Hove: Psychology Press. 71–114.
Rosen, C. 1990. Rethinking Southern Tiwa: the Geometry of a Triple-agreement Language. In *Language* 66. 669–713.
Ross, M. 1998. Possessive-like Attribute Constructions in the Oceanic Languages of Northwest Melanesia. In *Oceanic Linguistics* 37. 234–76.
Rottland, F. 1982. *Die Südnilotischen Sprachen. Beschreibung, Vergleichung und Rekonstruktion*. Kölner Beiträge zu Afrikanistik 7. Berlin: Dietrich Reimer.
Rouveret, A. and J.-R. Vergnaud. 1980. Specifying Reference to Subject: French Causatives and Conditions on Representation. In *Linguistic Inquiry* 11. 97–202.
Rude, N. 1997. Dative Shifting and Double Objects in Sahaptin. In T. Givón (ed.). *Grammatical Relations: A Functionalist Perspective*. Amsterdam: Benjamins. 323–49.
Rudzka-Ostyn, B. 1996. The Polish Dative. In W. van Belle and W. van Langendonck (eds.). *The Dative*. Amsterdam: John Benjamins. 341–94.
Ruhlen, M. 1987. *A Guide to the World's Languages*. Volume 1: *Classification*. London: Edward Arnold.
Ruigendijk, E. 2002. *Case Assignment in Agrammatism: A Cross-linguistic Study.* Doctoral dissertation. Groningen University, Groningen, The Netherlands.
—— and N. Friedmann. 2002. Not All Grammatical Morphemes are Impaired in Agrammatism: Evidence from the Production of Case in Hebrew and Dutch. In *Israeli Journal of Language, Speech and Hearing Disorders* 24. 41–54. (In Hebrew)
—— R. Van Zonneveld, and R. Bastiaanse. 1999. Case Assignment in Agrammatism. In *Journal of Speech, Language and Hearing Research* 42. 962–71.
—— and R. Bastiaanse R. 2002. Two Characteristics of Agrammatic Speech: Omission of Verbs and Omission of Determiners, Is There a Relation? In *Aphasiology* 16. 383–95.
—— N. Vasić, and S. Avrutin. 2006. Reference Assignment: Using Language Breakdown to Choose Between Theoretical Approaches. In *Brain and Language* 96. 302–17.
Rumpel, T. 1845. *Casuslehre in besonderer Beziehung auf die griechische Sprache*. Halle.
——1866. *Zur Casustheorie*. Programme de Gütersloh.
Russell, B. 1905. On Denoting. In *Mind* n.s. 14. 479–93.
Rutgers, R. 1998. *Yamphu: Grammar, Texts, and Lexicon*. Leiden: CNWS.
Ryabova, T. 1970. *Psikholingvisticheskii i neiropsikohologicheskii analiz dinamicheskoi afazii* [Psycholinguistic and neuropsychological analysis of dynamic aphasia]. Dissertation, Moscow.
Sadler, L. and R. Nordlinger. 2006. Case Stacking in Realizational Morphology. In *Linguistics* 44. 459–88.
Saffran, E. M., R. S. Berndt, and M. F. Schwartz. 1989. The Quantitative Analysis of Agrammatic Production: Procedure and Data. In *Brain and Language* 37. 440–79.

978–0–19–920647–6 59-Spencer-Ref OUP284-Malchukov-Spencer (Typeset by Spi, Delhi) 879 of 894 May 16, 2008 19:1

SAITO, M. 1992. Long Distance Scrambling in Japanese. In *Journal of East Asian Linguistics* 1. 69–118.

SAKSENA, A. 1980. The Affected Agent. In *Language* 56. 812–26.

SALTARELLI, M. 1988. (with M. Azkarate et al.) *Basque*. London: Routledge/Croom Helm.

SANDFELD, K. 1930. *Linguistique balkanique*. Paris: Libraire ancienne Honoré Champion.

SANDØY, H. 1996. *Talemål*. Oslo: Novus.

SAPIR, E. 1917. Review of *Het passieve karakter van het verbum transitivum of van het verbum actionis in talen van Noord-America* by C. C. Uhlenbeck. In *IJAL* 1. 82–6.

——1921. *Language: An Introduction to the Study of Speech*. Orlando/San Diego/New York/London: Harcourt Brace Jovanovich.

——1922. The Takelma Language of South-Western Oregon. In F. Boas (ed.). *Handbook of American Indian Languages*, Part 2. New York: Smithsonian Institution. 1–295.

SASSE, H.-J. 1984. Case in Cushitic, Semitic and Berber. In J. Bynon (ed.). *Current Progress in Afro-Asiatic Linguistics*. Amsterdam/Philadelphia: Jon Benjamins. 111–26.

SAVELJEVA, L. V. 1965. *Jazyk gudzharati* [The Gujarati Language]. Moscow: Nauka.

SAVICKIENĖ, I. 2002. The Emergence of Case Distinctions in Lithuanian. In M. D. Voeikova and W. U. Dressler (eds.). *Pre- and Protomorphology: Early Phases of Morphological Development in Nouns and Verbs*. Lincom Studies in Theoretical Linguistics 29. Munich: Lincom Europa. 105–15.

——2003. *The Acquisition of Lithuanian Noun Morphology*. Vienna: Verlag der Österreicheschen Akademie der Wissenschaften.

SAWADA, H. 2002. *Ninchi kategorii to zokusei jojutsu ni kansuru nichū taishō kenkyū* [Studies on cognitive categories and property predications in Japanese and Chinese]. PhD thesis, Kobe University.

SCATTON, E. 1984. *A Reference Grammar of Modern Bulgarian*. Columbus: Slavica Publishers.

——1993. Bulgarian. In B. Comrie and G. G. Corbett (eds.). *The Slavonic Languages*. London: Routledge. 188–248.

SCHACHTER, P. 1976. The Subject in Philippine Languages: Topic, Actor, Actor-Topic, or None of the Above? In C. N. Li (ed.). *Subject and Topic*. New York: Academic Press. 491–518.

——1977. Reference-Related and Role-Related Properties of Subjects. In P. Cole and J. Sadock (eds.). *Syntax and Semantics 8: Grammatical Relations*. New York: Academic Press. 279–306.

SCHADEBERG, T. C. 1986. Tone cases in Umbundu. In *Africana Linguistica* 10. 423–47.

——1990. *A Sketch of Umbundu*. Cologne: Rüdiger Köppe.

SCHEBECK, B. 1976. Thangu and Atjnjamathanha. In R. M. W. Dixon (ed.). *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies. 516–50.

SCHEEPERS, C. and M. M. B. CORLEY. 2000. Syntactic Priming in German Sentence Production. In L. R. Gleitman and A. K. Joshi (eds.). *Proceedings of the Twenty-Second Meeting of the Cognitive Science Society.* Mahwah, NJ: Erlbaum. 435–40.

——B. HEMFORTH, and L. KONIECZNY. 2000. Linking Syntactic Functions with Thematic Roles: Psych-Verbs and the Resolution of Subject–Object Ambiguity. In B. Hemforth and L. Konieczny (eds.). *German Sentence Processing*. Dordrecht: Kluwer. 95–135.

SCHIEFFELIN, B. B. 1985. The Acquisition of Kaluli. In D. I. Slobin (ed.). *The Crosslinguistic Study of Language Acquisition*. Vol. 1. Hillsdale, NJ: Erlbaum. 525–93.

Schiffman, H. F. 1983. *A Reference Grammar of Spoken Kannada*. Seattle: University of Washington Press.

Schlesewsky, M., G. Fanselow, R. Kliegl, and J. Krems. 2000. The Subject Preference in the Processing of Locally Ambiguous Wh-Questions in German. In B. Hemforth and L. Konieczny (eds.). *German Sentence Processing*. Dordrecht: Kluwer. 65–93.

Schlesinger, I. M. 1979. Cognitive Structure and Semantic Deep Structures: The Case of the Instrumental. In *Journal of Linguistics* 15. 203–395.

——1982. *Steps to Language: Toward a Theory of Native Language Acquisition*. Hillsdale, NJ: Erlbaum.

Schmitt, R. (ed.). 1989. *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert.

Schneider-Blum, G. 2003. *Case System in Alaaba*. Paper presented at the University of Cologne. Ms.

Schnur, T. T., A. Costa, and A. Caramazza. 2006. Planning at the Phonological Level during Sentence Production. In *Journal of Psycholinguistic Research* 35. 189–213.

Schriefers, H., A. D. Friederici, and K. Kühn. 1995. The Processing of Locally Ambiguous Relative Clauses in German. In *Journal of Memory and Language* 34. 499–520.

——E. Teruel, and R. M. Meinhausen. 1998. Producing Simple Sentences: Results from Picture–Word Interference Experiments. In *Journal of Memory and Language* 39. 609–32.

Schröder, H. 2002. *Word Order in Toposa. As Aspect of Multiple Feature-checking*. Nairobi. Ms.

Schuetze, C. T. 1996. Evidence for Case-related Functional Projections in Early German. In J. Camacho, L. Choueiri, and M. Watanabe (eds.). *Proceedings of the 14th West Coast Conference on Formal Linguistics (WCCFL)*. Stanford, CA: CSLI Publications. 447–61.

——and K. Wexler. 1996. Subject Case Licensing and English Root Infinitives. In A. Stringfellow, D. Cahana-Amitay, E. Hughes, and A. Zukowski (eds.). *Proceedings of the 20th Annual Boston University Conference on Language Development*. Vol. 2. Somerville, MA: Cascadilla Press. 670–81.

Schultze-Berndt, E. and N. P. Himmelmann. 2004. Depictive Secondary Predicates in Crosslinguistic Perspective. In *Linguistic Typology* 8. 59–131.

Schulze, W. n.d. *Functional Grammar of Udi.* Ms.

Schwarz, O., C. Stroh, and A. Urdze. 2001. *BIB-KOM-TYP. Die Bibliographie zur Komitativ-Typologie*. Munich: Lincom Europa.

Scott, G. 1978. *The Fore Language of Papua New Guinea*. Canberra: Pacific Linguistics.

Seefranz-Montag, A.v. 1983. *Syntaktische Funktionen und Wortstellungsveränderung: Die Entwicklung 'subjektloser' Konstruktionen in einigen Sprachen*. Munich: Wilhelm Fink.

Seiler, H. 1974. The Principle of Concomitance: Instrumental, Comitative and Collective (with special reference to German). In *Foundations of Language* 12. 215–47.

——1977. *Cahuilla Grammar*. Banning, CA: Malki Museum Press.

——1983. *Possession as an Operational Dimension of Language*. Tübingen: Gunter Narr Verlag.

Sells, P. 1998. The Function of Voice Markers in the Philippine Languages. In S. Lapointe et al. (eds.). *Morphology and Its Relation to Phonology and Syntax*. Stanford, CA: CSLI Publications. 111–37.

——2000. Raising and the Order of Clausal Constituents in the Philippine Languages. In I. Paul, V. Phillips, and L. Travis (eds.). *Formal Issues in Austronesian Linguistics*. Dordrecht: Kluwer Academic Publishers. 117–43.

——2001. Form and Function in the Typology of Grammatical Voice Systems. In G. Legendre, J. Grimshaw, and S. Vikner (eds.). *Optimality-Theoretic Syntax*. Cambridge, MA: MIT Press. 355–91.

——2004. Oblique Case Marking on Core Arguments. In J.-B. Kim, and B.-S. Park (eds.). *Perspectives on Korean Case and Case Marking*. Seoul: Thaehaksa. 151–82.

Serbat, G. 1981. *Cas et fonctions*. Paris: Presses Universitaires de France.

Seržants, I. 2004. Einige Bemerkungen zur Geschichte des Illativs. In *Baltu Filoloģija* 13 (1). 113–20.

Serzisko, F. 1982. Numerus/Genus-Kongruenz und das Phänomen der Polarität am Beispiel einiger ostkuschitischer Sprachen. In H. Seiler and F. J. Stachowiak (eds.). *Apprehension: das sprachliche Erfassen von Gegenständen* (vol. 2). Tübingen: Narr. 179–200.

——1992. *Sprechhandlungen und Pausen: Diskursorientierte Sprachbeschreibung am Beispiel des Ik*. Linguistische Arbeiten 282. Tübingen: Max Niemeyer.

Sgall, P., E. Hajičová, and J. Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects* (ed. by Jacob L. Mey). Dordrecht and Boston: Reidel.

Shapiro, J. N. and I. A. Chistovich. 2000. *Rukovodstvo po ocenke urovnja razvitija detej ot 1;2 do 3;6 mesjacev po rusificirovannoj škale RCDI-2000.* [Handbook of the estimation of the level of development of children from 1;2 to 3;6 with the help of the scale RCDI-2000 adapted to Russian]. Saint Petersburg: Institute of Early Intervention.

Sharma, D. 2001. Kashmiri Case Clitics and Person Hierarchy Effects. In P. Sells (ed.). *Formal and Empirical Issues in Optimality Theoretic Syntax*. Stanford, CA: CSLI Publications. 225–56.

Sharp, J. 2004. *Nyangumarta: A Language of the Pilbara Region of Western Australia*. Canberra: Pacific Linguistics.

Shay, E. and U. Seibert (eds.). 2003. *Motion, Direction and Location in Languages*. Amsterdam/Philadelphia: John Benjamins.

Shibatani, M. 1977. Grammatical Relations and Surface Cases. In *Language* 53. 789–809.

——1988. Voice in Philippine Languages. In M. Shibatani (ed.). *Passive and Voice*. Amsterdam: John Benjamins. 85–142.

——1990. *The Languages of Japan*. Cambridge: Cambridge University Press.

——1994. An Integrational Approach to Possessor Raising, Ethical Datives, and Adversative Passives. In *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society*. 461–86.

——2001. Non-canonical Constructions in Japanese. In A. Aikhenvald, R. W. M. Dixon, and M. Onishi (eds.). *Non-canonical Marking of Subjects and Objects*. (Typological studies in languages, 46). Amsterdam: Benjamins. 307–55.

——(ed.). 2002. *The Grammar of Causation and Interpersonal Manipulation*. Amsterdam: John Benjamins.

——2006. On the Conceptual Framework for Voice Phenomena. In *Linguistics* 44–2. 217–69.

——Z. Qín, and L. Tāo. 1994. Chinese Benefactive Constructions: Toward a Formal Analysis of the Schema-based Cognitive Approach. In M. Chen and O. Tzeng (eds.). *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*. Taipei: Pyramid Press. 459–77.

Siewierska, A. 1988. *Word Order Rules*. London: Croom Helm.

——1996. Word Order Type and Alignment Type. In *Sprachtypologie und Universalienforschung* 49. 149–76.

Siewierska, A. 1998. Variation in Major Constituent Order: A Global and a European Perspective. In A. Siewierska (ed.). *Constituent Order in the Languages of Europe*. Berlin: Mouton de Gruyter. 475–551.

—— 2003. Person Agreement and the Determination of Alignment. In *Transactions of the Philological Society* 101.2. 339–70.

—— 2004. *Person*. Cambridge: Cambridge University Press.

—— and D. Bakker. 1996. The Distribution of Subject and Object Agreement and Word Order Type. In *Studies in Language* 20–1. 115–62.

Sigurðsson, H. Á. 1991. Icelandic Case-marked PRO and the Licensing of Lexical Arguments. In *Natural Language and Linguistic Theory* 9. 327–64.

—— 1996. Icelandic Finite Verb Agreement. In *Working Papers in Scandinavian Syntax* 57. 1–46.

Silverstein, M. 1976. Hierarchy of Features and Ergativity. In R. M. W. Dixon (ed.). *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies. 112–71.

—— 1981. Case Marking and the Nature of Language. In *Australian Journal of Linguistics* 1. 227–46.

—— 1993. Of Nominatives and Datives: Universal Grammar from the Bottom up. In R. D. Van Valin (ed.). *Advances in Role and Reference Grammar*. Amsterdam: John Benjamins. 465–98.

Simpson, J. 1991. *Warlpiri Morpho-Syntax: A Lexicalist Approach*. Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer.

Sims-Williams, N. 1982. The Double System of Nominal Inflection in Sogdian. In *Transactions of the Philological Society*. 67–76.

Skjærvø, P. O. 1989a. Pashto. In R. Schmitt (ed.). *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert. 384–410.

—— 1989b. Yidgha and Munǰī. In R. Schmitt (ed.). *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert. 411–16.

Skorik, P. J. 1961. *Grammatika čukotskogo jazyka*, vol. 1. [A grammar of Chukchi. Vol. 1.] Leningrad: Izdatel'stvo Akademii nauk SSSR.

Slobin, D. I. 1966. The Acquisition of Russian as a Native Language. In F. Smith and G. A. Miller (eds.). *The Genesis of Language: a Psycholinguistic Approach*. Cambridge, MA: MIT Press. 129–48.

—— 1973. Cognitive Prerequisites for the Development of Grammar. In C. A. Ferguson and D. I. Slobin (eds.). *Studies of Child Language Development*. New York: Holt, Reinhart and Winston, Inc. 175–208.

—— 1982. Universal and Particular in the Acquisition of Language. In L. Gleitman and E. Wanner (eds.). *Language Acquisition: The State of the Art*. Cambridge: Cambridge University Press. 128–72.

—— 1985. Crosslinguistic Evidence for the Language-making Capacity. In D. I. Slobin (ed.). *The Crosslinguistic Study of Language Acquisition*, vol. 2. Hillsdale, NJ: Erlbaum. 1157–256.

De Smedt, K. 1990. IPF: An Incremental Parallel Formulator. In R. Dale, C. Mellish, and M. Zock (eds.). *Current Research in Natural Language Generation*. London: Academic Press. 167–92.

Smith, I. R. 1978. Sri Lanka Creole Portuguese Phonology. In *International Journal of Dravidian Linguistics* 7 (2). 248–406.

Au: Please updation of page number.

Smith, K. D. 1979. *Sedang Grammar: Phonological and Syntactic Structure*. Pacific Linguistics B, 50. Canberra: Australian National University.

Smith, M. B. 1993. Cases as Conceptual Categories: Evidence from German. In R. A. Geiger, and B. Rydzka-Ostyn (eds.). *Conceptualizations and Mental Processing in Language*. Berlin/New York: Mouton de Gruyter. 531–65.

Smith, M. and L. Wheeldon. 1999. High Level Processing Scope in Spoken Sentence Production. In *Cognition* 73. 205–46.

Smith, S. and E. Bates. 1987. Accessibility of Case and Gender Contrasts for Agent–Object Assignment in Broca's Aphasics and Fluent Anomics. In *Brain and Language* 30. 8–32.

—— and I. Mimica. 1984. Agrammatism in a Case-inflected Language: Comprehension of Agent–Object Relations. In *Brain and Language* 21. 274–90.

Sobolev, A. N. 1991a. Kategorija padeža na periferii balkanoslavjanskogo areala. In *Zbornik Matice srpske za filologiju i lingvistiku*. Vol. XXXIV/1. Novi Sad. 93–139.

—— 1991b. K istorii utraty balkanoslavjanskogo sklonenija. In *Zbornik Matice srpske za filologiju i lingvistiku*. Vol. XXXIV/2. Novi Sad. 7–41.

—— 1998. *Sprachatlas Ostserbiens und Westbulgariens*. Vols. I–III. Marburg: Biblion Verlag.

—— 2001. *Bolgarskij širokolykskij govor. Sintaksis. Leksika duhovnoj kul'tury. Teksty*. Marburg. Biblion Verlag.

—— (ed.). 2005. *Malyj dialektologičeskij atlas balkanskih jazykov. Serija grammatičeskaja*. Vol. I. *Kategorii imeni suščestviteljnogo*. Munich: Biblion Verlag.

Sohn, H. 1994. *Korean*. London: Routledge.

Song, J. J. 2001. *Linguistic Typology: Morphology and Syntax*. London: Longman.

Spencer, A. 1991. *Morphological Theory*: *An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell Publishers.

—— 1992. Nominal Inflection and the Nature of Functional Categories. In *Journal of Linguistics* 28. 313–41.

—— 2000. Agreement Morphology in Chukotkan. In W. U. Dressler, O. E. Pfeiffer, and M. A. Pochtrager (eds.). *Morphological Analysis in Comparison.* Amsterdam: Benjamins. 191–222.

—— 2003. A Realizational Approach to Case. In M. Butt and T. H. King (eds.). *Online Proceedings of the LFG03 Conference.* Stanford: CSLI Publications. 387–401. (http://csli-publications.stanford.edu/LFG/8/lfg03.html)

—— 2005. Case in Hindi. In M. Butt and T. H. King (eds.). *Proceedings of LFG '05*. Stanford: CSLI Publications. 429–46. http://csli-publications.stanford.edu/LFG/10/lfg05.html

—— 2006. Syntactic vs. Morphological Case: Implications for Morphosyntax. In L. Kulikov, A. Malchukov, and P. de Swart (eds.). *Case, Valency and Transitivity*. Amsterdam: John Benjamins. 3–21.

—— (forthcoming). Does Hungarian Have a Case System? In G. G. Corbett and M. Noonan (eds.). *Case and Grammatical Relations: Papers in Honour of Bernard Comrie*. Amsterdam: John Benjamins.

—— and R. Otoguro. 2005. Limits to Case – A Critical Survey of the Notion. In M. Amberber and H. de Hoop (eds.). *Competition and Variation in Natural Languages. The Case for Case.* Oxford: Elsevier. 119–45.

Stachovsky, M. and A. Menz. 1998. Yakut. In L. Johanson and E. Csató (eds.). *The Turkic Languages*. Routledge Language Family Descriptions Series. London: Routledge. 417–34.

Stallings, L. M., M. C. MacDonald, and P. G. O'Seaghdha. 1998. Phrasal Ordering Constraints in Sentence Production: Phrase Length and Verb Disposition in Heavy-NP Shift. In *Journal of Memory and Language* 39. 392–417.

Stampe, D. 1994. Extract from the summary of a discussion at LinguistList, dated 13 April 1994. (http://www.umich.edu/~archive/linguistics/linguist.list/volume.5/no. 451–500)

Starosta, S., A. K. Pawley, and L. A. Reid. 1982. The Evolution of Focus in Austronesian. In S. A. Wurm and L. Carrington (eds.). *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2. Canberra: Pacific Linguistics. 145–70.

Stassen, L. 2000. AND-languages and WITH-languages. In *Linguistic Typology* 4 (1). 1–54.

—— 2003. Noun Phrase Conjunction: The Coordinative and the Comitative Strategy. In F. Plank (ed.). *Noun Phrase Structure in the Languages of Europe*. Berlin: Mouton de Gruyter, 761–817.

Steele, S. 1978. Word Order Variation: A Typological Survey. In J. A. Greenberg (ed.). *Universals of Human Language*, vol. 4. Stanford: Stanford University Press. 585–623.

Stemberger, J. P. 1982. The Nature of Segments in the Lexicon: Evidence from Speech Errors. In *Lingua* 56. 43–67.

Stephany, U. and M. Voeikova (eds.). 2007. *Development of Nominal Inflection in First Language Acquisition: a Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter.

Sternefeld, W. 1985. Deutsch ohne grammatische Funktionen: Ein Beitrag zur Rektions- und Bindungstheorie. In *Linguistische Berichte* 99. 394–439.

Stevanović, M. 1975. *Savremeni srpskohrvatski jezik: gramatički sistemi i književnojezička norma*. 3rd edn. Belgrade: Zavod za udčbenike i nastavna sredstava.

Stiebels, B. 2000. Linker Inventories, Linking Splits and Lexical Economy. In B. Stiebels and D. Wunderlich (eds.). *Lexicon in Focus*. Berlin: Akademie Verlag. 211–245.

—— 2002. *Typologie des Argumentlinkings: Ökonomie und Expressivität*. Berlin: Akademie.

Stilo, D. L. 2001. Gilaki. In E. Yarshater (ed.). *Encyclopaedia Iranica*, Volume X. New York: Bibliotheca Persica Press. 660–8.

—— 2003. Hamadan: Jewish Dialect. In E. Yarshater (ed.). *Encyclopaedia Iranica,* Volume XI. New York: Bibliotheca Persica Press. 623–7.

—— 2004. *Vafsi Folk Tales*. Wiesbaden: Reichert.

—— 2006. Circumpositions as an Areal Response: The Case Study of the Iranian Zone. In L. Johanson and C. Bulut (eds.). *Turkic–Iranian Contact Areas: Historical and Linguistic Aspects*. Wiesbaden: Harrassowitz Verlag.

—— (2007a). Isfahan: Jewish Dialect. In E. Yarshater (ed.). *Encyclopaedia Iranica*, Volume XIV. New York: Bibliotheca Persica Press. 77–84.

—— (2007b). Gazi. In E. Yarshater (ed.). *Encyclopaedia Iranica*, Volume XIV. New York: Bibliotheca Persica Press. 112–19.

Stojkov, S. 1968. Nominale Kasusformen in der bulgarischen Sprache. In *Die Welt der Slaven* XIII. Wiesbaden. 28–46.

von Stockert, T. R. and L. Bader. 1976. Some Relations of Grammar and Lexicon in Aphasia. In *Cortex* 12. 49–60.

Stolz, T. 1992. *Lokalkasussysteme–Aspekte einer strukturellen Dynamik*. Wilhelmsfeld: Gottfried Egert.

—— 1996a. Komitativ-Typologie: MIT- und OHNE-Relationen im crosslinguistischen Überblick. *Papiere zur Linguistik* 51, 1: 3–65.

——1996b. Some Instruments are Really Good Companions – Some are Not: On Syncretism and the Typology of Instrumentals and Comitatives. In *Theoretical Linguistics* 23. 113–200.
——1997. Two Comitatives or More. On the Degree of Involvement of Participants. In A. Gather, and H. Werner (eds.). *Semiotische Prozesse und natürliche Sprache*. Stuttgart: Steiner. 515–30.
——1998. Komitative sind mehr als Instrumentale sind mehr als Komitative. Merkmalhaftigkeit und Markiertheit in der Typologie der MIT-Relationen. In K. Teržan-Kopecky (ed.). *Sammelband des 2. Internationalen Symposions zur Natürlichkeitstheorie*. Maribor: Universita v Mariboru. 83–100.
——2001a. Comitatives vs. Instrumentals vs. Agents. In W. Bisang (ed.). *Aspects of Typology*. Studio Typological. Berlin: Akademie Verlag. 153–74.
——2001b. To Be With X is To Have X: Comitatives, Instrumentals, Locative and Predicative Possession. In *Linguistics* 39/2. 321–50.
——2001c. On Circum-Baltic Instrumentals and Comitatives. To and Fro Coherence. In Ö. Dahl and M. Koptjevskaja-Tamm (eds.). *Circum-Baltic Languages*: *Typology and Contact*, Vol. 2: *Grammar and Typology*. Studies in Language and Companion Series 55. Amsterdam/Philadelphia: John Benjamins. 591–612.
——2003. Comitativo e strumentale nelle lingue germaniche e romanze. In L. Gaeta and S. Luraghi (eds.). *Introduzione alla linguistica cognitiva*. Rome: Carocci. 213–31.
—— C. Stroh, and A. Urdze. 2005. Comitatives and Instrumentals. In M. Haspelmath et al. (eds.). *The World Atlas of Language Structures*. Oxford: Oxford University Press. 214–17.
—— —— —— (forthcoming). *On Comitatives and Related Categories. A Typological Study with Special Focus on the Languages of Europe*. Berlin/New York: Mouton de Gruyter.
Stowe, L. A. 1986. Parsing Wh-Constructions: Evidence for On-Line Gap Location. In *Language and Cognitive Processes* 1. 227–45.
Stowell, T. 1981. *The Origins of Phrase Structure*. PhD Dissertation, MIT, Cambridge, MA.
——1982. The Tense of Infinitives. In *Linguistic Inquiry* 13. 561–70.
Strunk, J. 2004. *Possessive Constructions in Modern Low Saxon*. M.A. thesis, Stanford University.
Stump, G. T. 1993. On Rules of Referral. In *Language* 69. 449–79.
——2001a. *Inflectional Morphology. A Theory of Paradigm Structure.* Cambridge: Cambridge University Press.
——2001b. Affix Position. In M. Haspelmath, E. König, W. Oesterreicher, and W. Raible (eds.). *Language Typology and Language Universals*, volume I. Berlin: de Gruyter. 708–14.
Sturt, P., M. Pickering, and M. W. Crocker. 1999. Structural Change and Reanalysis Difficulty in Language Comprehension. In *Journal of Memory and Language* 40. 136–50.
Sumbatova, N. R. and R. O. Mutalov. 2003. *A Grammar of Icari Dargwa*. Munich/Newcastle: Lincom Europa.
Sussex, R. 1993. Slavonic Languages in Emigration. In B. Comrie and G. C. Corbett (eds.). *The Slavonic Languages*. London: Routledge. 999–1037.
Svenonius, P. 2002. Icelandic Case and the Structure of Events. In *Journal of Comparative Germanic Linguistics* 5. 197–225.
Svorou, S. 1994. *The Grammar of Space*. Amsterdam/Philadelphia: John Benjamins.

DE SWART, P. 2003. *The Case Mirror*. MA-thesis. Radboud University, Nijmegen.
—— 2006. Case Markedness. In L. Kulikov, A. Malchukov, and P. de Swart (eds.). *Case, Valency and Transitivity*. Amsterdam and Philadelphia: John Benjamins. 249–68.
—— (in prep.). PhD Dissertation, Radboud University, Nijmegen.
SWINNEY, D. and E. ZURIF. 1995. Syntactic Processing in Aphasia. In *Brain and Language* 50. 225–39.
SZEMERÉNYI, O. J. 1990. *Introduction to Indo-European Linguistics*. Oxford: Oxford University Press.
—— 1996. *Introduction to Indo-European Linguistics*. Oxford: Clarendon Press.
TAKAHASHI, C. 1994. Case, Agreement, and Multiple Subjects: Subjectivization in Syntax and LF. In N. Akatsuka (ed.). *Japanese/Korean Linguistics* 4. Stanford, CA: CSLI. 394–411.
TALIBOV, B. B. 1979. Morfologičeskaja i sintaksičeskaja xarakteristika padežej caxurskogo jazyka. In U. A. Mejlanova (ed.). *Imennoe sklonenie v dagestanskix jazykax*. Maxačkala: IIJaL.
TALMY, L. 2000. *Towards a Cognitive Semantics* (2 vols.). Cambridge, MA: MIT Press.
TAMM, A. 2005. *Relations between Estonian Verbs, Aspect, and Case*. Doctoral Dissertation, Eötvös Loránd University, Budapest.
TAMURA, S. 2000. *The Ainu Language*. Tokyo: Sanseido.
TANENHAUS, M. K. and J. C. TRUESWELL. 1995. Sentence Comprehension. In J. Miller and P. D. Eimas (eds.). *Speech, Language, and Communication*. San Diego, CA: Academic Press. 217–62.
TANNEN, D. 1989. *Talking Voices*. Cambridge: Cambridge University Press.
TANZ, C. 1974. Cognitive Principles Underlying Children's Errors in Pronominal Case-marking. In *Journal of Child Language* 1. 271–6.
TAYLOR, J. R. 1996. *Possessives in English: An Exploration in Cognitive Grammar*. Oxford: Oxford University Press.
TEDESCO, P. 1926. Ostiranische Nominalflexion. In *Zeitschrift für Indologie und Iranistik* 4. 94–166.
TERESHKIN, N. I. 1961. *Očerki dialektov xantyjskogo jazyka.* Vol. 1. Moscow: Nauka.
TERRILL, A. 2003. *A Grammar of Lavukaleve*. Mouton Grammar Library 30. Berlin/New York: Mouton de Gruyter.
TESNIÈRE, L. 1959. *Éléments de Syntaxe Structurale*. Paris: Klincksieck.
TESTELETS, J. 1998. On the Two Parameters of Transitivity. In L. Kulikov and H. Vater (eds.). *Typology of Verbal Categories. Papers Presented to V. Nedjalkov on the Occasion of his 70th Birthday*. Tübingen: Niemeyer. 29–45.
THOMASON, S. G. 2001. *Language Contact*. Edinburgh University Press.
—— and T. KAUFMAN. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley/Los Angeles/London: University of California Press.
THOMPSON, L. C. 1987. *A Vietnamese Grammar* (Second edition). Honolulu: University of Hawaii Press.
THRÁINSSON, H. 1979. *On Complementation in Icelandic.* New York: Garland Press.
TIFFOU, E. 1999. *Parlons Bourouchaski*. Paris: L'Harmattan.
TIMBERLAKE, A. 1974. *The Nominative Object in Slavic, Baltic, and West Finnish*. Slavistiche Beiträge, Band 82. Munich: Otto Sagner.
TODAEVA, B. X. 1997. Baoan'skij jazyk. In *Jazyki mira: mongol'skie jazyki, tunguso-man'čžurskie jazyki, japonskij jazyk, korejskij jazyk*. Moscow: Indrik. 29–36.

Tomasello, M. 1992. *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.

—— 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge, MA: Cambridge University Press.

Tomić, M. 1984. Govor Sviničana. In *Srpski dijalektološki zborni*. Vol. XXX. 7–265.

Tosco, M. 1994. On Case Marking in the Ethiopian Language Area (with special reference to the subject marking in East Cushitic). In V. Brugnatelli (ed.). *Sem, Cam, Iafet Atti della 7a Giornata di Studi Camito-Semitica e Indeuropei*. Milan: Centro studi camito-semitici. 225–44.

—— 2001. *The Dhaasanac Language: Grammar, Texts, Vocabulary of a Cushitic Language of Ethiopia*. Cushitic Language Studies 17. Cologne: Rüdiger Köppe.

Trask, R. 1979. On the Origin of Ergativity. In F. Planck (ed.). *Ergativity; Toward a Theory of Grammatical Relations*. New York: Academic Press. 385–404.

Traugott, E. C. and R. Dasher. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Travis, C. 2002. La Metalengua Semántica Natural: The Natural Semantic Metalanguage of Spanish. In C. Goddard and A. Wierzbicka (eds.). *Meaning and Universal Grammar: Theory and Empirical Findings*, vol. I. Amsterdam: John Benjamins. 173–242.

Traxler, M. J. and M. Pickering. 1996. Case-Marking in the Parsing of Complement Sentences: Evidence from Eye-Movements. In *Quarterly Journal of Experimental Psychology* 49A. 991–1004.

Trueswell, J. C., M. K. Tanenhaus, and C. Kello. 1993. Verb-Specific Constraints in Sentence Processing: Separating Effects of Lexical Preference from Garden-Paths. In *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19. 528–53.

Tsunoda, T. 1981a. *The Djaru Language of Kimberley, Western Australia*. Canberra: Pacific Linguistics.

—— 1981b. Split Case-marking in Verb Types and Tense/Aspect/Mood. In *Linguistics* 19. 389–438.

—— 1985. Remarks on Transitivity. In *Journal of Linguistics* 21: 385–96.

—— 1988. Antipassives in Warrungu and other Australian Languages. In M. Shibatani (ed.). *Passive and Voice*. Amsterdam: John Benjamins. 595–649.

Tsvetkova, L. S. 1969. *O narushenii aktivnykh form ustnoi rechi pri dinamicheskoi afazii* [Disturbance of the Active forms of spoken speech in dynamic aphasia]. Voprosy Psikhologii 1.

—— and J. M. Glozman. 1975. A Neurolinguistic Analysis of Expressive Agrammatism in Different Forms of Aphasia. In *Linguistics* 154/155. 61–76.

Tucker, A. N. and M. A. Bryan. 1966. *Linguistic Analyses. The Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press.

—— and J. T. O. Mpaayei. 1955. *A Maasai Grammar*. London/New York/Toronto: Longmans, Green and Co.

Tuomikoski, R. 1978. Objektinsijaisista määrän adverbiaaleista [Measure adverbials in the object cases]. In *Virittäjä* 82. 26–51.

Učida, N. 1970. *Der Bengali-Dialekt von Chittagong. Grammatik, Texte, Wörterbuch*. Wiesbaden: Harrassowitz

Ultan, R. 1978. Towards a Typology of Substantival Possession. In J. Greenberg et al. (eds.). *Universals of Human Language*, vol. 4: *Syntax*. Stanford: Stanford University Press. 11–50.

Uuspõld, E. and A. Valmet. 2001. *Populjarnaja grammatika èstonskogo jazyka*. Tallinn: Valgus.

Vainikka, A. 1989. *Deriving Syntactic Representations in Finnish*. PhD dissertation, University of Massachusetts, Amherst.

—— and J. Maling. 1996. Is Partitive Case Inherent or Structural? In J. Hoeksema (ed.). *Partitives. Studies on the Syntax and Semantics of Partitive and Related Constructions*. Berlin: Mouton de Gruyter. 179–208.

Valenzuela, P. 1997. *Basic Verb Types and Argument Structures in Shipibo-Conibo.* PhD thesis, University of Oregon.

Van Valin, R. D. 1977. *Aspects of Lakhota Syntax*. Doctoral dissertation, University of California, Berkeley.

—— 1990. Semantic Parameters of Split Intransitivity. In *Language* 66(2). 221–60.

—— 1995. Toward a Functionalist Account of So-called Extraction Constraints. In B. Devriendt, L. Goossens, and J. van der Auwera (eds.). *Complex Structures: A Functionalist Perspective*. Berlin: Mouton de Gruyter. 29–60.

—— 1998. The Acquisition of WH-questions and the Mechanisms of Language Acquisition. In M. Tomasello (ed.). *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*. Mahwah, NJ: LEA. 221–49. Available on RRG web site: http://linguistics.buffalo.edu/research/rrg.html

—— 1999. Generalized Semantic Roles and the Syntax–Semantics Interface. In F. Corblin, C. Dobrovie-Sorin, and J.-M. Marandin (eds.). *Empirical Issues in Formal Syntax and Semantics* 2. The Hague: Thesus. 373–89. Available on RRG web site: http://linguistics.buffalo.edu/research/rrg.html

—— 2004. Semantic Macroroles in Role and Reference Grammar. In R. Kailuweit and M. Hummel (eds.). *Semantische Rollen*. Tübingen: Narr. 62–82. Available on RRG web site.

—— 2005. *Exploring the Syntax–Semantics Interface*. Cambridge: Cambridge University Press.

—— and R. LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.

—— and D. P. Wilkins. 1997. The Case for 'Effector': Case Roles, Agents, and Agency Revisited. In M. Shibatani and S. Thompson (eds.). *Grammatical Constructions: Their Form and Meaning. A Festschrift for Charles Fillmore*. Oxford: Oxford University Press.

Vallduvi, E. and E. Engdahl. 1996. The Linguistic Realization of Information Packaging. In *Linguistics* 34. 469–519.

Vasić, N., S. Avrutin, and E. Ruigendijk. 2006. Interpretation of Pronouns in VP-ellipsis Constructions in Dutch Broca's and Wernicke's Aphasia. In *Brain and Language* 96. 191–206.

Veenker, W. 1967. *Die Frage des finnougrischen Substrats in der russischen Sprache*. Bloomington: Indiana University Press.

Vendler, Z. 1967. *Linguistics in Philosophy*. Ithaca: Cornell University Press.

Verma, M. K. and K. P. Mohanan (eds.). 1990. *Experiencer Subjects in South Asian Languages*. Stanford, CA: CSLI Publications.

—— —— (eds.). 1991. *Experiencer Subjects in South Asian Languages*. Menlo Park: Stanford University Press.

Vermeer, W. 1991. The Mysterious North Russian Nominative Singular Ending *-e* and the Problem of the Reflex of Proto-Indo-European **-os* in Slavic. In *Welt der Slaven* 36. 271–95.

Vermeulen, R. 2005. Possessive and Adjunct Multiple Nominative Constructions in Japanese. In *Lingua* 115. 1329–63.

Vigliocco, G. and T. Zilli. 1999. Syntactic Accuracy in Sentence Production: The Case of Gender Disagreement in Italian Language-impaired and Unimpaired Speakers. In *Journal of Psycholinguistic Research* 28. 623–48.

Viks, Ü. 1992. *Morphological Dictionary of Estonian: Introduction and Grammar*. Tallinn: Institute of Language and Literature.

De Vincenzi, M. 1991. *Syntactic Parsing Strategies in Italien*. Dodrecht: Kluwer.

van de Visser, M. 2006. *The Marked Status of Ergativity*. PhD dissertation, Utrecht University.

Voeikova, M. D. 2002. The acquisition of Case in Typologically Different Languages. In M. D. Voeikova and W. U. Dressler (eds.). *Pre- and Protomorphology: Early Phases of Morphological development in Nouns and Verbs*. Lincom Studies in Theoretical Linguistics 29. Munich: Lincom Europa. 25–45.

—— and W. U. Dressler (eds.). 2002. *Pre- and Protomorphology: Early Phases of Morphological Development in Nouns and Verbs*. Lincom Studies in Theoretical Linguistics 29. Munich: Lincom Europa.

—— and N. Gagarina. 2002. Early Syntax, First Lexicon and the Acquisition of Case Forms by two Russian Children. In M. D. Voeikova and W. U. Dressler (eds.). *Pre- and Protomorphology: Early Phases of Morphological Development in Nouns and Verbs*. Lincom Studies in Theoretical Linguistics 29. Munich: Lincom Europa. 115–33.

Vogt, H. 1971. *Grammaire de la langue Géorgienne.* [A Grammar of Georgian]. Olso: Universitetsforlaget.

van der Voort, H. 2000. *A Grammar of Kwaza*. PhD dissertation. University of Leiden.

—— 2004. *A Grammar of Kwaza*. Mouton Grammar Library 29. Berlin: Mouton de Gruyter.

Vyxuxolev, V. V. 1964. *Singal'skij jazyk*. Moscow: Nauka.

Wackernagel, J. 1922. *Vorlesungen über Syntax*. Band 2. Basel: Birckhäuser.

Waelchli, B. 2005. *Co-compounds and Natural Coordination*. Oxford: Oxford University Press.

Wälchli, B. and F. Zúñiga. 2006. Source–Goal (In)difference and the Typology of Motion Events in the Clause. In *Sprachtypologie und Universalienforschung* 59-3. 284–303.

Wali, K. and O. Koul. 1997. *Kashmiri: A Cognitive-Descriptive Grammar*. London: Routledge.

Wassenaar, M. E. D. 2005. *Agrammatic Comprehension: An Electrophysiological Approach*. Defended 29-09-2005, RU; prepared: MPI. Prom. prof. dr. P. Hagoort (MPI series in psycholinguistics 31).

Watahomigie, L., J. Bender, P. Watahomogie, Sr., and A. Y. Yamamoto. 2001. *Hualapai Reference Grammar*. Osaka: Endangered Languages of the Pacific Rim.

Watters, D. E. 2002. *A Grammar of Kham*. Cambridge: Cambridge University Press.

Weber, D. J. 1989. *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley: University of California Press.

Weerman, F. 1989. *The V2 Conspiracy; A Syntactic and a Diachronic Analysis*. Dordrecht: Foris.

—— and J. Evers-Vermeul. 2002. Pronouns and Case. In *Lingua* 112. 301–38.

Wegener, H. 1990. Der Dativ - ein Struktureller Kasus. In G. Fanselow and S. W. Felix (eds.). *Strukturen und Merkmale Syntaktischer Kategorien*. Tübingen: Narr.

WEINREICH, U. 1953. *Languages in Contact*. London/The Hague/Paris: Mouton.

—— 1980. *On Semantics*. Philadelphia: University of Pennsylvania Press.

WENINGER, S. 1999. *Ge'ez*. Munich: Lincom.

WENTZEL, T. W. 1980. *Die Zigeunersprache (Nordrussischer Dialekt)*. Leipzig: VEB Verlag.

WHALEY, L. 1997. *Introduction to Typology: The Unity and Diversity of Languages*. Thousand Oakes: Sage Publications.

WHEELDON, L. R. and M. C. SMITH. 2003. Phrase Structure Priming: A Short Lived Effect. In *Language and Cognitive Processes* 18. 431–42.

WHITE, L. 2003. *Second Language Acquisition and Universal Grammar*. Cambridge, MA: Cambridge University Press.

WHITNEY, W. D. 1879. *Sanskrit Grammar. Including both the classical language and the older dialects of Veda and Brahmana*. [Reprint of 2nd edition]. Cambridge, MA: Harvard University Press.

—— 1889. *Sanskrit Grammar*. Cambridge, MA: Harvard University Press.

WICHMANN, S. 1996. The Degrammaticalization of Agentivity in Tlapanec. In E. Engberg-Pedersen et al. (eds.). *Content, Expression, and Structure. Studies in Danish Functional Grammar.* Amsterdam: Benjamins. 343–60.

—— (in press). The Reference-Tracking System of Tlapanec: Between Obviation and Switch Reference. In *Studies in Language.*

—— (forthcoming). Sandhi tonal interno en la morfología verbal tlapaneca. In *The Proceedings of the VIII Encuentro Internacional de Lingüística en el Noroeste*. Universidad de Sonora, Hermosillo, Mexico, Nov. 17–19, 2004.

WIERZBICKA, A. 1979. Ethno-syntax and the philosophy of grammar. In *Studies in Language* 3(3). 313–83.

—— 1980. *The Case for Surface Case*. Ann Arbor: Karoma.

—— 1981. Case Marking and Human Nature. *Australian Journal of Linguistics* 1. 43–80.

—— 1986a. The Meaning of a Case: A Study of the Polish Dative. In R. D. B. and J. S. Levine (eds.). *Case in Slavic*. Columbus, OH: Slavica Publishers.

—— 1986b. The Semantics of 'Internal Dative' in English. In *Quaderni di Semantica* 7(1). 121–35.

—— 1988. *The Semantics of Grammar*. Amsterdam: John Benjamins.

—— 1996. *Semantics: Primes and Universals*. Oxford: Oxford University Press.

—— 2002. Semantic Primes and Universal Grammar in Polish. In C. Goddard and A. Wierzbicka (eds.). *Meaning and Universal Grammar: Theory and Empirical Findings*. vol. 2. Amsterdam: John Benjamins. 65–144.

—— in press. Bodies and their Parts: an NSM Approach to Semantic Typology. In *Language Sciences*.

WILKINS, D. 1989. *Mparntwe Arrernte (Aranda): Studies in the Structure and Semantics of Grammar*. Canberra: The Australian National University.

WILLIAMS, E. 1980. Predication. In *Linguistic Inquiry* 11. 203–38.

WILSON, S. R. 1996. Verb Agreement and Case Marking in Burushaski. In *Working Papers of the Summer Institute of Linguistics, University of North Dakota*. 40. 1–71.

VAN WINDEKENS, A. J. 1979. *Le tokharien confronté avec les autres langues indo-européennes*, Vol. II/1: *La morphologie nominale*. Leuven: Peeters.

WINDFUHR, G. L. 1992. Cases. In E. Yarshater (ed.). *Encyclopaedia Iranica*, Volume V. New York: Bibliotheca Persica Press. 25–37.

WOLDEMARIAM, H. 2003. *The Grammar of Haro with Comparative Notes on the Ometo Linguistic Group*. PhD dissertation, University of Addis Ababa.

WOODBURY, A. 1977. Greenlandic Eskimo, Ergativity, and Relational Grammar. In P. Cole and J. Sadock (eds.). *Syntax and Semantics 8: Grammatical Relations*. News York: Academic Press. 307–36.

WOOLFORD, E. 1995. Object Agreement in Palauan: Specificity, Humanness, Economy and Optimality. In J. N. Beckman, L. W. Dickey, and S. Urbanczyk (eds.). *Papers in Optimality Theory*. University of Massachusetts Occasional Papers, 18: Amherst: GLSA. 655–700.

—— 2001. Case Patterns. In G. Legendre, J. Grimshaw, and S. Vikner (eds.). *Optimality-theoretic Syntax*. Cambridge, MA: MIT Press. 509–45.

—— 2006. Lexical Case, Inherent Case, and Argument Structure. In *Linguistic Inquiry* 37.1. 111–30.

—— 2008. Differential Subject Marking at Argument Structure, Syntax, and PF. In H. de Hoop and P. de Swart (eds.). *Differential Subject Marking* Dordrecht: Springer. 17–40.

WÜLLNER, F. 1831. *Über Ursprung und Urbedeutung der Sprachlichen Formen*. Münster.

WUNDERLICH, D. 1997. Case and the Structure of Verbs. In *Linguistic Inquiry* 28. 27–68.

—— 2006. Argument Hierarchy and Other Factors Determining Argument Realization. In I. Bornkessel, M. Schlesewsky, B. Comrie, and A. D. Friederici (eds.). *Semantic Role Universals and Argument Linking. Theoretical, Typological, and Psycholinguistic Perspectives*. Berlin: de Gruyter. 15–52.

—— and R. LAKÄMPER. 2001. On the Interaction of Structural and Semantic Case. In *Lingua* 111. 377–418.

WURM, S. A. 1976. Accusative Marking in Duungidjawu (Waga-Waga). In R. M. W. Dixon (ed.). *Grammatical Categories in Australian Languages*. New Jersey: Humanities Press. 106–11.

WURMBRAND, S. 2001. *Infinitives: Restructuring and Clause Structure.* Berlin/New York: Mouton de Gruyter.

—— 2005. *Tense in Infinitives*. Plenary talk at the workshop on *New Horizons in the Grammar of Raising and Control*. LSA Summer Institute, Cambridge, MA.

—— 2006. Licensing Case. In *Journal of Germanic Linguistics* 18.3. 175–234.

—— (To appear). Infinitives are Tenseless. In *Proceedings of the 30th Annual Penn Linguistics Colloquium: Penn Working Papers in Linguistics* 13.1.

WURZEL, W. U. 1984. *Flexionsmorphologie und Natürlichkeit: ein Beitrag zur morphologischen Theoriebildung*. Berlin (East): Akademie-Verlag.

VAN DEN WYNGAERD, G. 1989. Object Shift as an A-Movement Rule. In *MIT Working Papers in Linguistics* 11. 256–71.

XRAKOVSKIJ, V. S. and A. P. VOLODIN. 1986. *The Semantics and Typology of the Imperative: The Russian Imperative*. Leningrad: Nauka.

YADROFF, M. 1996. Modern Russian Vocatives: A Case of Subtractive Morphology. In *Journal of Slavic Linguistics* 4. 133–53.

YALLOP, C. 1977. *Alyawarra: An Aboriginal Language of Central Australia*. Canberra: AIAS Press.

YAMAGUCHI, K. 2004. *A Typological, Historical, and Functional Study of Adpositions in the Languages of the World*. PhD thesis, University of New Mexico.

YAMANASHI, M. 1995. *Ninchi bunpō ron* [Studies on cognitive grammar]. Tokyo: Hitsuji-Shobō.

Yamashita, H. and F. Chang. 2006. Sentence Production in Japanese. In M. Nakayama, R. Mazuka, and Y. Shirai (eds.). *Handbook of East Asian psycholinguistics*, Vol. 2: *Japanese*. Cambridge: Cambridge University Press.

Yarshater, E. 1969. *Southern Tati Dialects*. The Hague/Paris: Mouton.

—— 1970. The Tati dialect of Ṭārom. In M. Boyce and I. Gershevitch (eds.). *W. B. Henning Memorial Volume*. London: Lund Humphries. 451–67.

Yip, M. 1998. Identity Avoidance in Phonology and Morphology. In S. G. Lapointe, D. K. Brentari, and P. M. Farrell (eds.). *Morphology and Its Relation to Phonology and Syntax*. Stanford: CSLI Publications. 216–46.

—— J. Maling, and R. Jackendoff. 1987. Case in Tiers. In *Language* 63. 217–50.

Yoon, K-J. 2006. *Constructing a Korean Natural Semantic Metalanguage*. Seoul: Hankook Publishing Company.

—— 2007. ADD. In M. Amberber (ed.). *The Language of Memory in a Crosslinguistic Perspective*. Amsterdam/Philadelphia: John Benjamins.

Zaenen, A., J. Maling, and H. Thráinsson. 1985. Case and Grammatical Functions: The Icelandic Passive. In *Natural Language and Linguistic Theory* 3. 441–83. [Reprinted in J. Maling and A. Zaenen (eds.). *Syntax and Semantics 24: Modern Icelandic Syntax*. New York: Academic Press. 95–164.]

Zaliznjak, A. A. 1967. *Russkoe imennoe slovoizmenenie* [Russian nominal inflection]. Moscow: Nauka.

—— 1973 [2002]. O ponimanii termina 'padež' v lingvističeskix opisanijax. In A. A. Zaliznjak (ed.). *Problemy grammatičeskogo modelirovanija*. Moscow: Nauka. 53–87. [Reprinted in *Russkoe imennoe slovoizmenenie*. Moscow: Jazyki slavjanskoj kul'tury. 613–47.

Zeevat, H. and G. Jäger. 2002. A Reinterpretation of Syntactic Alignment. In D. de Jongh, H. Zeevat, and M. Nilsenova (eds.). *Proceedings of the Third and Fourth International Tbilisi Symposium on Language, Logic and Computation*. Amsterdam: University of Amsterdam.

Zekokh, U. S. 1969. *Sklonenie v adygejskom jazyke*. Majkop: Adygejskoe otdelenie Krasnodarskogo knižnogo izdatel'stva.

Zhukova, A. N. 1972. *Grammatika korjakskogo jazyka*. [Koryak Grammar]. Leningrad: Nauka.

Zinkevičius, Z. 1996. *The History of the Lithuanian Language*. Vilnius: Mokslo ir enciklopedijų leidykla.

Žirkov, L. I. 1955. *Lakskij jazyk*. Moscow: Izdatel'stvo Akademii Nauk SSSR.

Zograf, G. A. 1976. *Morfologičeskij stroj novyx indoarijskix jazykov: opyt strukturno-tipologičeskogo analiza* [The morphological structure of the New Indo-Aryan languages: a structural-typological approach]. Moscow: Nauka.

Zovko, I. 2000. *Dative Shift in Croatian*. Paper presented at the Conference on New Theoretical Approaches to Syntax and Semantics in Cognitive Science, Dubrovnik.

—— 2001. *SemanticËko-sintakticËki odnosi u recËenicama s dvostruko prijelaznim glagolima u engleskom jeziku*. M.A. thesis, University of Zagreb.

Zúñiga, F. 2006. *Deixis and Alignment: Inverse Systems in Indigenous Languages of the Americas*. Amsterdam: Benjamins.

Zwicky, A. 1974. Hey, Whatsyourname! In M. W. LaGaly, R. A. Fox, and A. Bruck (eds.). *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 787–801.

——1985. How to Describe Inflection. In M. Niepokuj et al. (eds.). *Proceedings of BLS*, 11. Berkeley Linguistics Society. 372–86.

——1992. Jottings on Adpositions, Case Inflections, Government, and Agreement. In D. Brentari, G. Larson, and L. MacLeod (eds.). *The Joy of Grammar: A Festschrift for James D. McCawley*. Amsterdam: John Benjamins. 369–83.

——and G. K. Pullum. 1983. Cliticization versus Inflection: English *n't*. In *Language* 59. 502–13.

—— ——1987. Plain Morphology and Expressive Morphology. In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: University of California at Berkeley. 330–40.